Studies of U.S. Navy Cues, Information Order, and Impact of Conflicting Data

M. J. Liebhaber
D. A. Kobus
Pacific Science and Engineering Group, Inc.

B. A. Feher
SSC San Diego

SSC San Diego

20021008 178

TECHNICAL REPORT 1888
August 2002

Studies of U.S. Navy Cues, Information Order, and Impact of Conflicting Data

M. J. Liebhaber
D. A. Kobus
Pacific Science and Engineering Group, Inc.

B. A. Feher
SSC San Diego

SPAWAR
Systems Center
San Diego

SSC SAN DIEGO
San Diego, California 92152-5001

T. V. Flynn, CAPT, USN
Commanding Officer

R. C. Kolb
Executive Director

ADMINISTRATIVE INFORMATION

ACKNOWLEDGMENTS

SB

EXECUTIVE SUMMARY

Three studies investigated the threat assessment process for shipboard air defense (AD). The goal was to better understand the relationship between air track information (e.g., altitude, speed, country of origin) available to a ship's AD personnel and their perceived level of threat regarding a particular aircraft. Understanding this process is crucial in designing effective AD decision support tools.

BACKGROUND

No one system or person on the ship does not provide aircraft identification (ID) or threat assessment. These tasks result from interaction among sensor systems, computers, and human operators. Highly trained personnel must evaluate, integrate, and judge information. The studies reported here started in the Tactical Decision-Making Under Stress (TADMUS) program. TADMUS evaluated Combat Information Center (CIC) Decision Support System (DSS) display concepts derived from cognitive theory. Displays supported the cognitive strategies of tactical decision-makers. Several studies evaluated the decision-making performance of AD personnel as they interacted with a dynamic AD scenario[1] (Kelly and Moore, 1996; Morrison, 2000).

The DSS included a threat assessment window that displayed a binary threat assessment (Threat or No Threat) and the underlying data (e.g., speed, range, etc.) used to generate the assessment. The window design was based on the concept of explanation-based reasoning in which decision-makers form a hypothesis and fit supporting and contradicting evidence into a plausible story (Pennington and Hastie, 1988 and 1993). However, theoretical and applied investigation of threat assessment concepts was minimal. Assessment and categorization of data were based on suggestions from experts, with the caution that future research would be conducted. A preliminary study[2] stressed the need for continued research. The study discovered elements about the threat assessment process that did not conform to assumptions of subject matter experts: Threat level is not always related to aircraft platform type or ID (e.g., Commercial/Nonmilitary [COMAIR]). AD personnel do not equally consider or weigh all available data.

FINDINGS

Participants in the three studies averaged 2.95 years of at-sea experience in one or more positions on the AD staff. Many participants were currently serving in those positions.

The goal of the first study was to detail the relationship among specific values of cues (i.e., cues, characteristics, or factors used to evaluate tracks such as Altitude, Speed, or Identification Friend or Foe [IFF] Mode) and their corresponding perceived threat ratings in littoral and open-ocean environments. Participants provided ratings of threat level and threat level change for 18 cues relevant to the threat assessment process. Results provided baseline threat levels for aircraft and a detailed list of the relationship among specific values of cues and the corresponding perceived threat ratings.

---

[1] S. G. Hutchins and D. P. Westra. 1996. "Tactical Decision-Making Under Stress: Baseline Results." Contact Bela Feher, SSC San Diego, CA.

[2] M. J. Liebhaber and C. A. P. Smith. 1999. "Basis for Assessment." Contact Bela Feher, SSC San Diego, CA.

The next study empirically evaluated the relative importance of the threat assessment cues to determine how AD personnel used the cues. The findings indicated that participants did not frequently use all cues. Participants consistently used only 6 to 13 cues out 18 available cues. They used fewer cues to evaluate Friendly tracks compared to Enemy tracks, and fewer cues to evaluate tracks in an open ocean than in a littoral environment. Six critical cues were identified based on their frequency and sequence of selection by participants. They were, in order of weight or importance, Origin, IFF Mode, Intelligence, Altitude, Airlane, and Electronic Signal (ES).

The final study assessed the effect of conflicting information on track assessment. Participants were given a track ID and all 18 cues. Some cues would occasionally contain data that conflicted with the ID. Conflicting data, when present, came from a track with a different ID. Results indicated that all but minimal amounts of conflicting data interfered with threat assessment. Participants were most likely to switch track ID when there was conflicting, high weighted data: Origin, IFF Mode, Intelligence, Altitude, and Airlane. However, knowing which cues were interfered with was more important than knowing the quantity of conflicting data. Participants changed track ID and threat levels because of conflicting data in the Origin, IFF Mode, and ES cues.

RECOMMENDATIONS

The experienced personnel in these studies did not choose to evaluate all available cues and they relied on some cues more than other cues. Decision Support Systems and displays need to consider these behaviors and their potential to introduce biases. A complete threat assessment model should evaluate all cues, display the threat level and contribution of each cue to the overall threat level, and provide input to a system that suggests an appropriate course of action. Understanding this process will lead to better guidelines for tactical situation displays. Effective displays will present information consistent with the threat assessment task and the user's mental model of that task.

CONTENTS

Figures

Tables

1. INTRODUCTION

This report discusses the results from a series of three studies that investigated the air threat assessment process. Air threat assessment is part of ship air defense (AD) and evaluates aircraft that are flying near one's ship or the battle group and determines the threat. These studies focused on how experienced U.S. Navy AD personnel conduct single-ship AD defense threat assessment.

Threat assessment is not a task assigned to any one person. It is an implicit process that underlies AD decision-making and occurs at many levels of operational responsibility within a ship's Combat Information Center (CIC). Evaluating air threats is a cognitively intense, time-critical process subject to human bias and error (Chalmers, 1998; Klein, 1993; Schulze et al., 1999; Van Sickle, 2000; Zimmerman, 1997). A thorough understanding of the process is crucial in understanding CIC situation assessment and decision-making. It is also crucial in creating informed guidelines for effective decision support systems.

Effectively assessing threat requires AD personnel to cognitively integrate information from many different sources. Integration requires a high level of tactical expertise, including knowledge of the types of threats, context (e.g., ship's mission and regional geopolitical climate), familiarity with U.S. Navy doctrine (e.g., Rules of Engagement, Operational Task Orders), and assessment heuristics built from experience. The multitasking nature of the AD task also introduces cognitive complexity (Chalmers, 1998). AD personnel must concentrate on one of several competing stimuli yet divide attention among multiple competing tasks. For example, the AD Commander must decide on a course of action about a particular aircraft while keeping aware of available resources (e.g., intercept and tanker aircraft, ship's weapons systems), monitoring video displays and audio broadcasts, and preparing periodic situation reports for the Commanding Officer.

Besides cognitive demands, proposed changes in the AD task environment may make the job even more complex. For example, the U.S. Navy has suggested that future ships operate with fewer personnel. Complex operations (e.g., delivering humanitarian aid in a war zone) and time pressure add stress to the AD task. In this environment, team members may find it difficult to notice or identify important pieces of information that may increase understanding of the tactical situation. Current displays are inadequate in conveying information to U.S. Navy decision-makers because data are spread across multiple systems and displays. Gathering and mentally integrating the data are especially difficult in highly dynamic situations (Chalmers, 1998; Chalmers, Easter, and Potter, 2000). Empirically derived and validated models and tools for decision support are needed. A better understanding of the threat assessment process is required to meet these needs.

2. BACKGROUND

2.1 TADMUS AND RELATED WORK

Decision support for CIC tasks, especially aircraft identification, has been studied from behavioral (Chalmers, 1998), organizational (Zimmerman, 1997), and engineering[1] perspectives. The studies started in the Tactical Decision-Making Under Stress (TADMUS) program (Morrison, 2000). TADMUS evaluated CIC Decision Support System (DSS) display concepts derived from cognitive theory. Displays supported the cognitive strategies of tactical decision- makers. Several studies evaluated the decision-making performance of AD personnel as they interacted with a dynamic AD scenario[2] (Kelly and Moore, 1996).

The DSS included a threat assessment window that displayed a binary threat assessment (Threat or No Threat) and the underlying data (e.g., speed, range, etc.) that were used to make the assessment. The window design was based on the concept of explanation-based reasoning in which decision makers form a hypothesis and fit supporting and contradicting evidence into a plausible story (Pennington and Hastie, 1988; 1993). Data were displayed in tables with three categories: Supporting Evidence, Counter Evidence, and Assumptions. However, theoretical and applied investigation of threat assessment concepts was minimal. Data were assessed and categorized based on suggestions from experts (with the caution that future research would be conducted).

2.2 AIR DEFENSE

To understand the identification and threat assessment process, one must realize that assessment is provided by intense interaction among sensor systems, computers, and human operators. One system does not provide the identification or assessment of a particular air track. Highly trained personnel must evaluate, correlate, and judge information (Zimmerman, 1997). The dynamic nature of air warfare, short decision times, the large number of aircraft, the high volume of information, and the need to integrate the information while conducting multiple tasks increases complexity (Bell and Lyon, 2000; Chalmers, 1998; Chalmers, Easter, and Potter, 2000; Van Sickle, 2000).

An automated combat information system processes and displays air tracks for AD personnel. The Tactical Action Officer (TAO) leads the team, usually as AD Commander (ADC). Some important AD personnel include the Anti-Air Warfare Coordinator (AAWC), Electronic Warfare Supervisor (EWS), and EW Operator (EWO). Appendix A lists AD roles.

As experienced AD personnel evaluate a track, they must continuously monitor displays and audio communication nets, evaluate and exchange the information, and decide on an appropriate course of action. Processing includes such tasks as verifying an aircraft's identity or other important pieces of data (e.g., What kind of weapons does it have?), physically locating the track on a geo-plot (map display), and determining what resources are necessary to deal with the track. Most physical data come from various sensor systems such as air search radars, IFF interrogation systems, and electronic

---

[1] R. Lyons. 2000. "Joint Tactical Terminal (JTT) Operational Analysis Report." Contact Bela Feher, SSC San Diego, CA.

[2] S. G. Hutchins and D. P. Westra. 1996. "Tactical Decision-Making Under Stress: Baseline Results." Contact Bela Feher, SSC San Diego, CA.

emission detection systems. Aircraft threat evaluation is an extremely time- and resource-consuming process (Van Sickle, 2000).

Current combat systems such as those on Aegis ships use kinematics (e.g., changes in range, course, speed, and altitude) and tactical data (e.g., type of radar) to analyze and ID aircraft contacts. The system operators use these and other factors (e.g., intelligence) to classify aircraft. Typically, the AD team relies on doctrine and historical experience to devise classification schemes before the ship deploys. The AD team must also purposefully integrate other information such as geography, intelligence reports, and country of origin during the assessment process to determine a course of action. This integration requires a high level of skill and experience that is poorly documented.

2.3 ASSUMPTIONS ABOUT AD PROCESS

Discussions with subject matter experts and research participants raised some issues about the process that previous data did not support. Because threat assessment occurs as a natural part of air track evaluation, two assumptions seem to negate the need to study threat assessment. Making these assumptions explicit is critical because it reveals the complexity of the assessment process and demonstrates the need for a thorough understanding of the process.

The first assumption is that the AD team does not assess the threat level of an aircraft because threat level is related to aircraft ID (e.g., Commercial or Tactical Aircraft), which can be mechanically derived from a predetermined set of factors. This assumption appears to be based on the observation that assessment simply involves recognition (i.e., ID and classification). However, Liebhaber and Smith[3] did not support the assumption that threat level is related to ID. They found that aircraft ID was not strongly correlated with threat level. Sometimes commercial aircraft were rated as more threatening and demanded a higher priority than military aircraft. The findings suggested that experienced AD personnel were tacitly evaluating different factors besides ID. Liebhaber and Smith[4] suggested that perceived threat level was related to how far a track cue deviated from expected behavior. Aircraft with high deviations from expected behavior were rated as more threatening than aircraft with low deviations from expected behavior.

A second assumption is that AD personnel equally consider or weigh all available data. The process of evaluation and ID can be handled by planned, predetermined criteria. However, previous research does not support this assumption. Liebhaber and Smith[5] found that the research participants used only a small set of available track cues to assess the threat of an aircraft. The selected cues depended on their initial perception of the aircraft being evaluated. Table 1 shows the cues and Appendix B defines them. Participants tended to use more factors than those found in conventional predetermined criteria. They used factors in ways that were inconsistent with previous assumptions. For example, range is assumed to figure prominently in threat assessment; however, participants rarely used it. Therefore, it appears that AD personnel do not use all cues in all situations or in ways that conform to prior assumptions.

---

[3] M. J. Liebhaber and C. A. P. Smith. 1999. "Tactical Decision-Making Under Stress: Baseline Results." Contac Bela Feher, SSC San Diego, CA.

[4] Liebhaber and Smith, *ibid.*

[5] Liebhaber and Smith, *ibid.*

Table 1. Cues used to evaluate high- and low-threat tracks (ordered by frequency of use).

Low-Threat Track	High-Threat Track
ES and platform type	ES and platform type
Altitude	Weapon type
Speed	Weapon envelope
Airlane	Coordinated activity
Coordinated activity	Range
Origin	Course
IFF	Altitude
Course	Closest Point of Approach (CPA)
	Origin
	IFF
	Feet wet
	Speed
	Intelligence
	Available support
	Airlane

3. CURRENT STUDIES

Experienced AD personnel rely on various factors to evaluate air tracks. Examples include altitude, speed, and country of origin. A series of three studies were conducted to better understand the relationship between the information available to the AD team and the level of threat that they perceive regarding a particular track. The first study was an extension of the Liebhaber and Smith[6] investigation of the cues and processes of threat assessment. The second experiment was an empirical evaluation of how qualified AD team members selected, grouped, and ranked specific cues. The third study identified the effect of confirming, missing, or conflicting data on a perceived threat. Because little is known about how cue data values influence threat level, the hypothesis tests for the studies were conducted with alpha equal to 0.1.

3.1 STUDY 1: CUE AND THREAT LEVEL RELATIONSHIP

This study details the relationship between specific values of each cue and the corresponding perceived threat ratings. Liebhaber and Smith[7] evaluated participants as they assessed the threat levels of tracks in a computer-simulated AD scenario. This investigation included a broader range of track types and track behaviors.

3.1.1 Participants

Sixteen U.S. Navy personnel with at-sea experience in AD participated in this study. Demographic information was collected from each participant. The participants had a mean of 3.3 years of at-sea CIC experience. Most had experience in more than one role on the AD team. Appendix A describes each role.

3.1.2 Design

A questionnaire was used to gather data on the relationship between track cues and perceived threat level. Participants were asked to respond to a series of questions related to two Areas of Operational Responsibility (AORs): the Northern Persian Gulf and the Northern Pacific Ocean. "Threatening" was defined as the perceived ability of a track to inflict damage on one's own ship or battle group. It pertained only to the degree of the threat posed by a particular track.

3.1.2.1 Air Track Analysis Questionnaire. Appendix C shows the questionnaire. The questionnaire clarifies the relationship between information that AD team members used to evaluate tracks and the level of threat that the team perceived. It asked three types of questions. The first set of questions (items 1, 2, and 3 in Appendix C) identify baseline threat levels for track ID classifications (e.g., Unknown, Neutral, Enemy, etc.), weapons (e.g., Harpoon, Exocet, etc.), and radar emitters.

In the next set of questions (items 4 through 22 in Appendix C), participants estimated change to the baseline threat level. Each item had a list of data values or ranges (Table 2). The ranges were determined from transcripts of verbal protocols in Liebhaber and Smith (1999) and from subject matter experts. Participants were instructed to treat each cue (e.g., speed) as if it were the first piece of data received, and then to indicate how the data would change their threat estimate. They were told

---

[6] Liebhaber and Smith, *ibid.*

[7] Liebhaber and Smith, *ibid.*

to answer based on their at-sea experience. All questions were about single pieces of information (e.g., speed, altitude, or IFF Mode). Although information is not processed in isolation in the real situation, our goal was to understand how individual elements were evaluated. Participants responded on a five-point Likert scale. For example, participants were asked, "If a track had a speed within the ranges shown below, how would it change your estimate of the threat posed by that track?"

Table 2. Estimated change to baseline threat level with data values or ranges.

Speed Ranges	Lower Greatly	Lower a Little	No Change	Raise a Little	Raise Greatly
Under 150 knots	1	2	3	4	5
150 to 250 knots	1	2	3	4	5
250 to 350 knots	1	2	3	4	5
350 to 450 knots	1	2	3	4	5
450 to 550 knots	1	2	3	4	5
Over 550 knots	1	2	3	4	5

The final questions (items 22, 23, and 24 in Appendix C) asked participants to identify from whom they received information, to whom they passed information, and what type of information was passed. The purpose of this question was to understand the flow of data and the functional relationships within CIC.

3.1.2.1.1 Independent Variable. Area of Operational Responsibility (AOR): Participants rated threat levels and threat changes for each item in the Northern Persian Gulf and the Northern Pacific Ocean AORs.

3.1.2.1.2 Dependent Variables. Data were collected on a five-point Likert scale (see Appendix C for scale labels).

- Baseline Threat Level for

  o Track classifications (e.g., Unknown, Neutral, Enemy, etc.)

  o Types of weapons

  o Radar emitters.

- Threat Change Rating (TCR) for each data range for each cue. TCR is the mean change in perceived threat level. It equals the grand mean response minus three (the "No change" response). For example, if a mean response of 1.8 is calculated for a specific cue, then TCR will be $1.8 - 3.0 = -1.2$. Thus, the TCR is the magnitude of change and the increase or decrease in baseline threat level.

3.1.3 Hypotheses

Naval operations in littoral settings (e.g., Northern Persian Gulf) are more complex and stressful than operations in the open ocean (Chalmers, 1998) because the near-shore environment is characterized by increased numbers of air and surface craft and decreased response time to potential threats. Therefore, one might expect that participants would assign higher threat levels to tracks in littoral settings than in the open ocean.

3.1.4 Procedure

Each participant was briefed about the purpose of the study. They filled out a written questionnaire that included questions about their military experience and track threat ratings. The questionnaire took about 45 minutes to finish.

3.1.5 Results

The baseline threat levels for categories of tracks and types of weapons and radar emitters were reported first, followed by threat level changes and information flow. The effect of AOR on threat ratings was analyzed.

3.1.5.1 Baseline Threat Ratings for Track ID Categories. Participants were asked to rate the perceived threat of tracks in the standard categories (e.g., Friend, Enemy, etc.). Table 3 lists mean ratings. Values ranged from Low Threat (1.0) to High Threat (5.0). Friendly tracks were consistently rated as less threatening than other ID types. For each ID category, ratings in the Northern Persian Gulf AOR were higher than those in the Open Ocean AOR. The values in Table 3 are considered the baseline or default threat levels for tracks in those ID categories.

Table 3. Mean Threat Level for each track ID category.

ID Category	Northern Persian Gulf	Northern Pacific Ocean
Friend	1.50	1.00
Assumed friend	2.50	2.00
Unknown	3.20	2.50
Assumed enemy	3.50	3.50
Hostile	4.50	4.00

3.1.5.2 Baseline Threat Ratings for Types of Aircraft Weapon Systems. Table 4 (N = 13) shows participants' ratings of the degree of threat posed by selected weapons systems. Contact the authors for data on other weapons systems. Values ranged from Low Threat (1.0) to High Threat (5.0).

Table 4. Mean Threat Level by AOR for each type of weapon system.

Weapon Type	Northern Persian Gulf	Northern Pacific Ocean
Exocet	4.5	3.5

9

3.1.5.3 Baseline Threat Ratings for Types of Radar Emitters. Participants assigned a threat level to several types of radar emitters. Additional emitter information was part of the ES cue, and is reported below. Appendix D shows the ES threat levels. Values ranged from Low Threat (1.0) to High Threat (5.0).

3.1.5.4 TCRs for Each Cue. Cue values and their corresponding TCRs were tabulated. Appendix D shows the computed TCRs for each cue in alphabetical order. Appendix B provides definitions of each cue. Figure 1 shows examples of threat level changes for aircraft speed. TCR tends to decrease as speed increases.
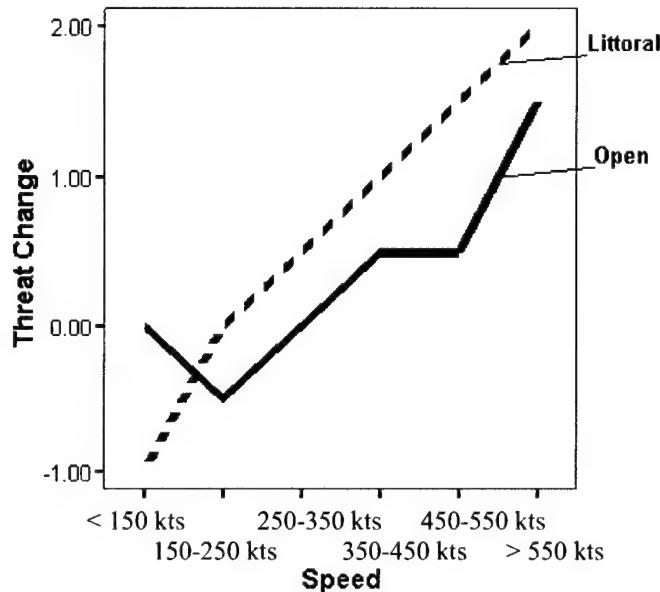
Figure 1. Change in threat level as a function of speed and AOR (0.0 = no change).

3.1.5.5 Effect of AOR. The effect of AOR (Northern Persian Gulf, Open Ocean) on threat levels and TCRs was evaluated using one-way analysis of the variance (ANOVA). Mean responses were computed for each cue range (e.g., Speed: 450 to 550 knots) across all participants. Parametric analysis is appropriate for summated rating scales where there is an underlying continuous distribution (Gaito, 1980; Michell, 1986; Velleman and Wilkinson, 1993). There were no significant differences in baseline threat levels for ID category or Weapon Type. TCRs for Airlane, Altitude, Communications, Confidence, Course (includes CPA), ES, Feet Wet/Dry, IFF Mode, Number (Composition), Origin, Range, Speed, Support, Wings Clean/Dirty, Warning Status, and Weapon Envelope were also analyzed. Significant differences at the $p < 0.1$ level were found for the following:

- Course (Steady and Closing)

- CPA (25 to 50 nmi)

- Feet Dry

- Feet Wet (perpendicular to and >5 nmi from coast)

- Voice Communication (None)

- Speed (450 to 550 knots)

3.1.5.6 CIC Information Flow. Participants were asked to identify from whom they received information and to whom they passed information. The goal was to obtain a better understanding of the flow of data and the functional relationships within CIC. Figure 2 shows the flow of information that participants reported. The shaded boxes indicate the roles participants held in this study. The lines in Figure 2 are double-coded with solid lines, indicating two-way communication and dashed lines denoting one-way communication. The arrowhead indicates the direction. Line thickness represents the number of times a participant mentioned a given connection; thicker lines mean higher frequency. Results indicate that the AAWC plays a crucial role in assimilating and transmitting information.
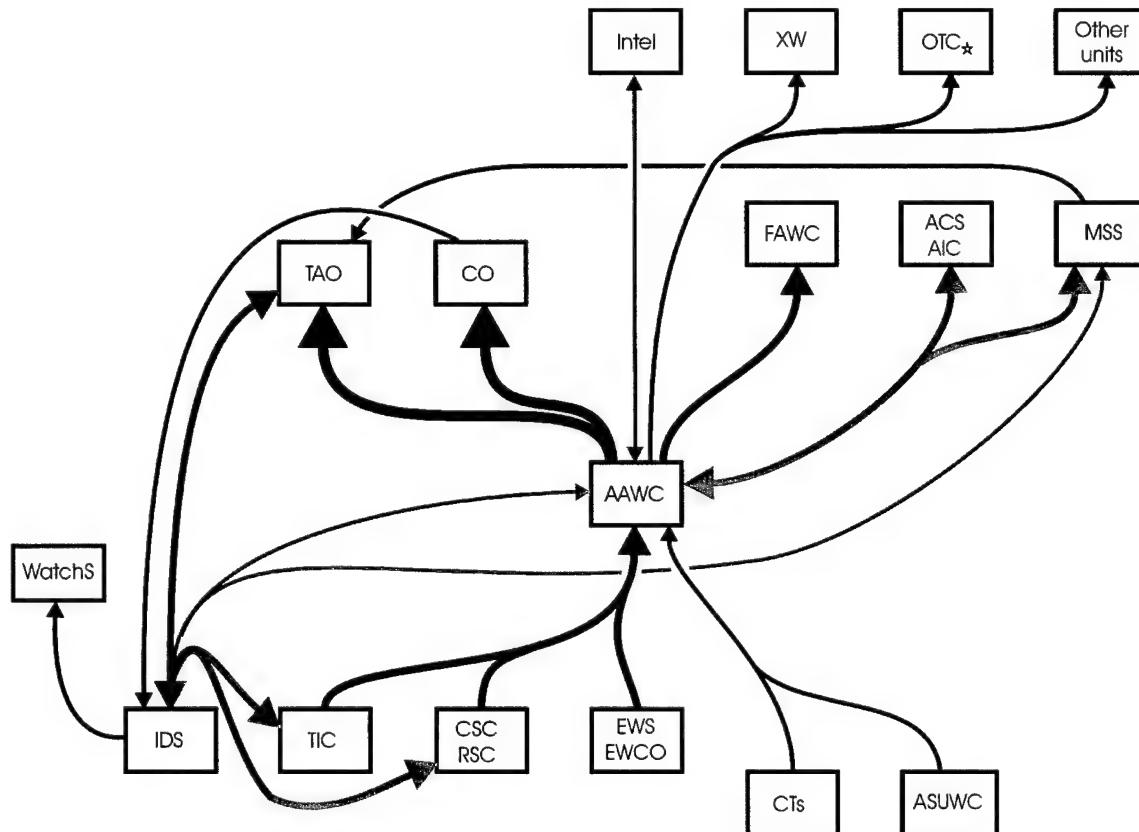
Figure 2. Reported information flow within CIC.

3.1.6 Study 1 Summary

Study 1 provided specifications of the relationship among track cues and perceived threat ratings. Participants provided ratings of threat level or threat level change for cues relevant to the threat assessment process. Some differences in threat change values were due to AOR. Most notable was a CPA of 25 to 50 nmi. Tracks with this CPA value were perceived as much less threatening in the Northern Pacific Ocean as compared to a Northern Persian Gulf setting. Not surprisingly, tracks in

11

the Friendly ID category were rated as less threatening than tracks in the Hostile ID category. Unknown tracks were rated as an intermediate threat. Tracks in the Northern Persian Gulf AOR setting were rated as more threatening than tracks in the Open Ocean AOR setting despite the ID category.

While this study described the relationship between track cues and changes to threat level, it did not identify how the cues were used to evaluate tracks, the relative importance of the cues, or if they were used in pairs or groups. The data from this study were then used to create the stimulus materials for Studies 2 and 3.

3.2 STUDY 2: RELATIVE IMPORTANCE OF THREAT ASSESSMENT CUES

Liebhaber and Smith (1999) found that participants consistently used some threat assessment cues more than other cues and that they appeared to evaluate the cues in a fairly regular sequence. Cues that were evaluated earlier and more often were assumed more important than later, less frequently selected cues. Important cues are cues that provide the most evidence or information value to the assessment process. Liebhaber and Smith[8] did not address cue importance directly. Therefore, this study empirically evaluated the importance of the threat assessment cues. Experienced AD personnel made ID and threat assessment decisions during simulated scenarios. These data were collected to develop guidelines for tactical situation displays. Effective displays provide information in a manner that is consistent with a task and the user's mental model of that task (Vicente and Rasmussen, 1992). If AD personnel process track data in terms of importance or in combination, displays that conform to those expectations are likely more supportive of knowledge-based decision-making than the current displays.

3.2.1 Participants

Eighteen U.S. Navy personnel participated in this study. They were serving in an AD position and averaged from 1.5 years to 4.8 years of at-sea AD experience. Most participants were experienced in more than one role. Appendix A defines the AD roles.

3.2.2 Design

A 2 (AOR) X 3 (Given ID) X 2 (Matching) within subjects design was used.

3.2.2.1 Independent Variables. The definitions of the variables are as follows:

- AOR (Northern Pacific Ocean, Northern Persian Gulf): Area of Operational Responsibility for a given trial. The Northern Pacific Ocean represents an open ocean environment and the Northern Persian Gulf is a littoral setting.

- Given ID (Assumed Friend, Unknown Evaluated, or Assumed Enemy): Track ID shown to the participant on the experimental display.

- Data Congruence (Yes or No): The track data that the participants saw were constructed so that the data either matched the Given ID Category (Congruence = Yes) or did not match the Given ID (Congruence = No).

---

[8] Liebhaber and Smith, *ibid.*

3.2.2.2 Dependent Variables. The dependent variables were as follows:

- Cue: The cue(s) that were selected on a given trial. Participants could select up to 18 cues on each trial.

- Cue Selection Time: Time interval (msec) between each cue a participant selected.

- ID Accuracy (Yes or No): Participants were required to select an ID category (Friend, Unknown, Enemy) based on the data they evaluated. Responses were scored as Correct when the ID that was assigned by the participant matched the Actual ID of the track on a given trial. Otherwise, responses were scored as Incorrect.

- ID Time: Time interval between the last selected cue and the ID choice.

- Threat Level: Threat level that participant identified. Ranged from 1 (Low) to 7 (High).

- Threat Time: Time interval between ID choice and threat level selection.

- N Cues: Number of cues that were selected by the subject.

3.2.3 Hypotheses

1. Cue Use:

   a. Participants will use fewer cues to evaluate Friend tracks than Enemy tracks.

   b. Participants will use fewer cues to evaluate tracks in the Northern Pacific Ocean than in the Northern Persian Gulf AOR.

   c. Participants will use certain cues more often than other cues.

2. Accuracy:

   a. Participants will accurately identify all categories (Friend, Unknown, Enemy) of tracks.

3. Threat Level:

   a. Friendly tracks will be perceived as less threatening than Enemy tracks.

   b. Tracks in the Northern Persian Gulf AOR will be perceived as more threatening than tracks in the Northern Pacific Ocean AOR.

4. Cue Selection Time:

   a. Participants will take longer to select cues for Unknown and Hostile tracks than for Friendly tracks.

   b. Participants will take longer to select cues in the Northern Persian Gulf AOR than in the Northern Pacific Ocean AOR.

3.2.4 Procedure

Volunteer AD personnel filled out a questionnaire about their experience and participated in a laptop-computer-based experiment. The participants' task was to confirm the ID shown on the screen (the Given ID) and then assign a level of threat to the track. They underwent six 5-minute sessions. The entire research session lasted about 45 to 50 minutes.

3.2.4.1 Display. Figure 3 shows the initial display for a typical trial. From top to bottom, the display components were as follows:

- AOR: Either Northern Pacific or North Persian Gulf.

- Track ID: The Given-ID (Assumed Friend, Unknown Evaluated, or Assumed Enemy).

- Time Remaining: Running count of time left. Participants were given 5 minutes to evaluate as many tracks as possible. The counter ran from 5 minutes down to 0 minutes.

- ID Accuracy: Continually updated percentage of times the participant's ID matched the actual ID of the track.

- Trial Number: Current trial number.

- Available Task Information: Labeled buttons for each cue. Participants could select any number of cues in any sequence. Cues were selected by using the mouse to click on the button. When participants clicked on a cue, a data value appeared (e.g., Altitude = 22,000 ft). Figure 4 shows how the display would look after the participant selected two cues. Participants were free to select cues that they thought were most appropriate for the given situation to resolve their assessment of threat.

- ID: Participants selected an ID by clicking on one category. Once selected, they could not select any more cues.

- Threat Level: Participants indicated a threat level by clicking on one of the seven buttons.

Figure 3. Sample screen for typical trial.
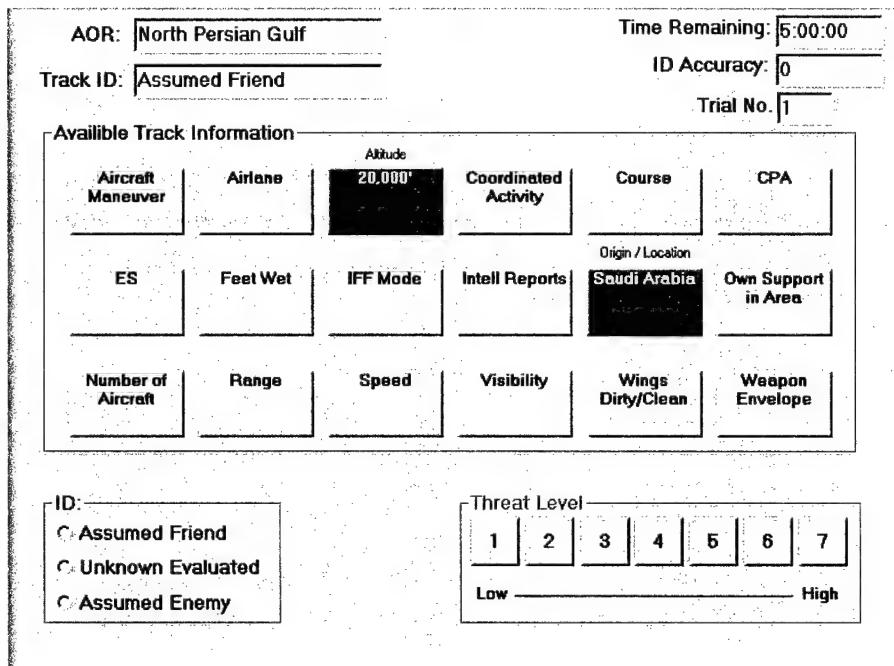
Figure 4. Sample screen after participant selected Origin and Altitude.

3.2.4.2 Sequence of Events

1. Participants were briefed about the task and given the following instructions to read. The instructions were available to the participant during the experiment.

Instructions

- Situation:

  a. You are the Anti-Air Warfare Coordinator (AAWC) onboard an Aegis-class cruiser.

  b. You have just come on duty and must assume the watch.

  c. You have:

  i. not received a "Pass Down" or other briefing, and

  ii. a series of air tracks to evaluate.

- ID Tasking:

  a. An initial Track ID has been suggested.

  b. Available supporting information is displayed.

  c. Country of Origin indicates point of initial detection.

- ADW Mission and Tasking

  a. Maintain a Unit Self-Defensive AAW Posture.

15

b. Conduct Positive Identification (PID) by evaluating and correlating available data.

c. Assess the level of threat posed by each track to own ship.

2. Participants were then shown the components of the display and were told the following:

a. Look at the AOR and Initial ID for each new experimental trial.

b. Click on the factors that you want to review.

c. Select an ID and Threat Level.

d. Once you select a Track ID, you cannot uncover more data.

e. After you select a Threat Level, the computer will automatically display the next track.

3. You will have 5 minutes to evaluate as many tracks as possible.

4. Please work as quickly and as accurately as possible.

5. Tracks 1 and 2 are practice samples.

6. Participants went through two practice trials. After the practice trials, participant's remaining questions were answered.

7. When the participants were ready, the experimental trials began.

3.2.4.3 Experiment Data File. The input data file contained the information that was displayed on the screen for each trial. There were six input files, three for the North Gulf AOR and three for the Northern Pacific AOR. Each file contained two practice trials and 60 experimental trials. The experimental trials were presented randomly. There were 20 Assumed Friend, 20 Unknown Evaluated, and 20 Assumed Enemy trials. The track data for each cue was constructed to be consistent or inconsistent with the Given ID in Table 5.

Table 5. Track data for each cue.

Given ID	Total	Matched Assumed Friend	Matched Unknown Evaluated	Matched Assumed Enemy
Assumed Friend	20	8	4	8
Unknown valuated	20	8	4	8
Assumed Enemy	20	8	4	8

3.2.5 Results

The results are reported by hypothesis. To ease interpretation, only trials where the Given ID Category matched the Actual ID (i.e., Matching = Yes) were analyzed.

3.2.5.1 Hypothesis 1a and 1b: Number of Cues Selected by Participants by AOR and Track ID.
Although all 18 cues were available for selection on each trial, participants only selected between 6 and 13 cues per a given trial (mean = 9.6, SD = 4.3). Significantly fewer cues were selected when the

data matched a Friendly profile as opposed to an Enemy or Unknown profiles (F (2, 434) = 39.7, p < .05). Participants also used fewer cues when they evaluated tracks in the Northern Pacific Ocean AOR, but only for Unknown and Enemy tracks. ID of the track (F (2, 434) = 139.9, p < 0.05). Figure 5 shows the interaction (N_Pac = Northern Pacific Ocean, N_Gulf = Northern Persian Gulf).
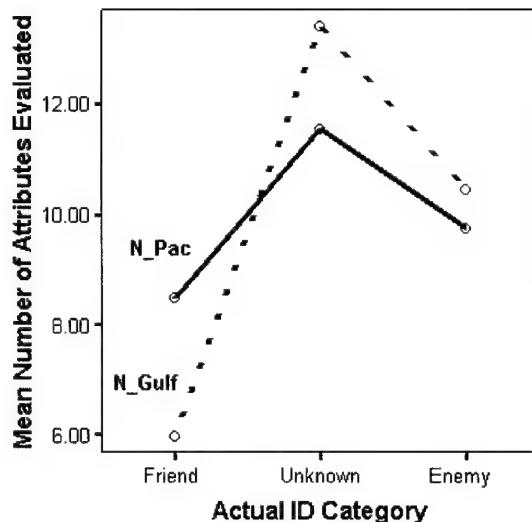
Figure 5. Mean number of cues selected per trial in each AOR.

3.2.5.2 Hypothesis 1c: Cue Selection Order. Table 6 shows the average order, across all trials, in which participants selected cues. Origin was most often selected first, followed by Intelligence, and so on. Frequency is the number of times each cue was chosen in the given sequential position. Origin was selected first more times (630) than any other cue, overall and regardless of AOR. Duplicates occurred because of bimodal selection distributions for some cues (e.g., in the Northern Persian Gulf AOR, Wings Dirty was chosen equally often in positions 10 and 12).

Table 6. Mean order of cue selection across all participants based on frequency.

	Cues by AOR			Frequency of Selection		
Selection Order	Northern Persian Gulf AOR	Northern Pacific Ocean AOR	Overall	Northern Gulf	North. Pacific Ocean	Overall
1	Origin	Origin	Origin	291	339	630
2	Intel	Intel	Intel	148	183	331
3	IFF	IFF	IFF	145	193	338
4	Airlane	Airlane	Airlane	98	104	202
5	ES	ES	ES	114	103	217
6	Maneuver	Maneuver	Maneuver	85	98	183
7	Number	CPA	Number	66	59	119
8	Feet wet	Airlane	Feet wet	61	47	105
9	Speed and IFF	Altitude	Altitude	42	51	91

17

Table 6. Mean order of cue selection across all participants based on frequency. (continued)

Selection Order	Cues by AOR			Frequency of Selection		
	Northern Persian Gulf AOR	Northern Pacific Ocean AOR	Overall	Northern Gulf	North. Pacific Ocean	Overall
10	Wings dirty	Speed	Speed	39	44	68
11	Speed	Speed	Speed	39	34	73
12	Wings dirty	Number & own support	Number & wings dirty	33	32	50
13	Speed & Altitude	CPA	CPA	19	29	44
14	Wings dirty	Wings dirty	Wings dirty	23	31	54
15	Visibility	Visibility	Visibility	24	25	49
16	Wings dirty	Wings dirty	Wings dirty	14	21	35
17	Weapon envelope	Range	Range	10	16	25
18	Weapon envelope	Number	Number	14	17	28

3.2.5.3 Hypothesis 1c: Relative Cue Weight. When the participants evaluated a track, they used from 6 to 13 cues and selected the cues in roughly the same sequence for each track. It was assumed that cues selected early and often were more important or critical to the threat assessment process than later and less frequently selected cues. However, selection order and frequency did not adequately capture the notion of cue importance, especially for cues such as Airlane and ES. When participants chose Airlane and ES, Airlane was usually chosen before ES. However, ES was selected more frequently than Airlane (see Table 6).

To quantify the concept of importance, an index of Relative Cue Weight (W) based on order and frequency was computed for each cue. The resulting index number represented the relative importance of each cue as a function of how often it was picked in a given order. Cues with a high index value were selected more frequently and sooner relative to the other cues.

A total Relative Cue Weight (W_{Cue}) was calculated for each cue. Participants could choose up to 18 cues, in any order, on each trial (selection position 1 = first cue selected by the participant, 2 = second selection, etc.). In general, a mean weight was computed for each selection position. Then W_{Cue} was calculated by summing across all selection positions (1 through 18). The formula for computing Relative Cue Weight (W) is as follows:

$$W_{Attribute} = \text{Sum across all selection positions} \left(\begin{array}{l} \text{Mean proportion that the current} \\ \text{Attribute was the selected in} \\ \text{position } i \text{ by all participants} \end{array} \quad * \quad \begin{array}{l} \text{Selection} \\ \text{Position } i\text{'s} \\ \text{Weight} \end{array} \right)$$

$$W_{Attribute} = \sum_{i=1}^{18} \left(\left(\frac{\sum_{j=1}^{N_{participants}} \frac{Frequency_j}{N_{Trials_j}}}{N_{Participants}} \right) * (19 - i) \right)$$

where,

        i (1 to 18) = Selection Position.

        j = Current participant ($N_{participants}$ = 18).

Mean proportion for selection position i = number of times the cue was chosen by participant j as selection number i, divided by the number of trials for participant j. Each participant's speed determined the number of trials they underwent in the 5-minute time limit. The mean proportion was computed by dividing the summed proportions by $N_{participants}$. This method allowed the variation in number of trials for each subject and did not give artificially high weight to responses from subjects who received a proportionally high number of trials. The proportion for each participant was divided by the total number of participants to give a mean proportion.

Selection position weight = each selection position was assigned a weight. The highest weight (18 = 19 - 1) was given to the cue that was selected first, then next highest (17 = 19 - 2) to the second cue selected, and so on.

Tables 7 and 8 show the computed Relative Cue Weights. The computed weights confirmed the prediction in Hypothesis 1c that participants were not relying equally on all cues. Based on the weights, the most critical cues were Origin, IFF Mode, Intelligence, Altitude, Airlane, and to some extent, ES.

Table 7. Relative Cue Weights (*W*) for each ID Category in Northern Pacific AOR.

Friend		Unknown		Enemy	
Cue	*W*	Cue	*W*	Cue	*W*
Origin	15.18	Origin	15.58	Origin	15.44
IFF mode	14.32	IFF mode	13.70	IFF mode	13.48
Intelligence	12.57	Intelligence	11.70	Intelligence	12.72
Altitude	10.68	Altitude	11.44	Airlane	11.57
Airlane	10.36	Airlane	10.93	Altitude	11.15
ES	8.41	ES	8.68	ES	9.26
CPA	6.80	CPA	7.57	Maneuvers	7.27
Speed	6.78	Speed	7.48	CPA	6.84
Feet wet	6.52	Coordinated act	7.08	Coord act	6.81
Coordinated act	6.20	Maneuvers	6.96	Feet wet	6.80
Number of tracks	6.20	Feet wet	6.55	Speed	6.75
Maneuvers	6.08	Number tracks	6.18	Number of tracks	6.47
Wings clean	4.09	Course	4.09	Course	4.70
Course	4.06	Range	4.04	Wings clean	4.39
Range	3.90	Wings clean	3.83	Range	4.18
Support	2.05	Support	1.61	Support	2.12
Visibility	1.12	Visibility	1.59	Visibility	1.16
Weapon envlp	1.05	Weapon envlp	0.84	Weapon envlp	0.63
Visibility	1.10	Visibility	1.48	Support	1.98
Weapon envlp	0.33	Weapon envlp	0.77	Weapon envlp	0.82

Table 8. Relative Cue Weights (*W*) for each ID Category in Northern Gulf AOR.

Friend		Unknown		Enemy	
Cue	*W*	Cue	*W*	Cue	*W*
Origin	15.48	Origin	15.55	Origin	15.84
IFF mode	12.32	IFF mode	14.04	IFF mode	14.07
Intelligence	10.67	Intelligence	11.87	Intelligence	12.11
Airlane	9.88	Airlane	11.61	Airlane	11.85
Altitude	9.77	Altitude	11.33	Altitude	11.46
ES	8.34	ES	9.04	ES	8.91
Speed	6.16	Feet wet	7.97	Feet wet	7.89
CPA	6.02	Coordinated act	7.39	Speed	7.65
Feet wet	6.01	CPA	7.38	CPA	7.36
Number of tracks	5.48	Speed	7.21	Coordinated act	6.83
Coordinated act	5.25	Maneuvers	6.31	Number of tracks	6.55
Maneuvers	5.04	Number of tracks	6.16	Maneuvers	5.83
Wings clean	3.61	Course	5.70	Course	5.18
Range	3.48	Range	4.21	Wings clean	4.73
Course	3.42	Wings clean	3.88	Range	4.54
Support	1.48	Support	2.23	Visibility	2.01
Visibility	1.10	Visibility	1.48	Support	1.98
Weapon envlp	0.33	Weapon envlp	0.77	Weapon envlp	0.82

In addition to preferring certain cues, participants appeared to be considering some cues in combination. Combinations were evident in the verbal protocols from our previous work, where certain cues tended to co-occur (e.g., Course and Speed). For the current analysis, the six highest weighted cues were divided into three groups based on naturally occurring breaks in *W* that can be observed in Tables 7 and 8. A hierarchical cluster analysis was also conducted on the weights from each ID Category (Friend, Unknown, and Enemy) to confirm the observed groupings. Three fairly stable clusters were observed across all IDs. They were as follows:

- Cluster 1:   Origin

                      IFF Mode

- Cluster 2:   Intelligence

                Altitude

                Airlane

- Cluster 3:   ES

A pair-wise mean difference analysis was conducted using the Fisher–Hayter Test (Kirk, 1995, p. 148) to confirm the differences in weights between the clusters. The relative weights in Cluster 1 were significantly higher than the weights in Cluster 2 for Origin versus Altitude (qFH (17,204) = 1.75, $p < 0.05$) and Origin versus Airlane (qFH (17,204) = 1.73, $p < 0.05$). The relative weights in Cluster 1 were also significantly higher than the weights in Cluster 3: Origin versus ES (qFH (17,204) = 2.60, $p < 0.05$) and IFF Mode versus ES qFH (17,204) = 1.88, $p < 0.05$. Clusters 2 and 3 were not significantly different from one another. Therefore, only two clusters could be confirmed.

5.2.5.4 Hypothesis 2: ID Accuracy. The number and proportion of correct responses are contained in boxes in Table 9. Participants correctly identified most of the tracks that they evaluated (*Pearson Chi-Square* (4) = 235.587, $p < 0.05$). They selected Friend on 66.9% of the trials when the Actual ID was also Friend (see Table 9). By comparison, participants responded Unknown or Enemy at a significantly lower rate. Participants were 67.1% accurate when Actual ID was Unknown and 70.5% accurate when Actual ID was Enemy. Accuracy was similar in both AORs: Northern Persian Gulf = 70.3% and Northern Pacific Ocean = 67.6%. It was also similar across all ID Categories: Friend = 66.9%, Unknown = 67.1%, and Enemy = 70.5%.

Table 9. Accuracy: Participant's ID Response compared to Actual ID.

	ID Response of Participant		
Actual ID	ID = Friend	ID = Unknown	ID = Enemy
Friend	89 (66.9%)	30 (22.6%)	14 (10.5%)
Unknown	3 (4.3%)	47 (67.1%)	20 (28.6%)
Enemy	26 (11.0%)	44 (18.6%)	167 (70.5%)

5.2.5.5 Hypothesis 2: ID Response Time. Although participants were accurate across all ID categories, they took more time to make an ID decision for Enemy tracks than for Friend tracks in both AORs (F (1.43, 22.887) = 7.347, $p = .007$, $\eta^2 = 0.315$). Figure 6 shows mean times. The sphericity assumption was violated in this analysis (Mauchly's $W = .602$, $p = 0.022$), therefore, the degrees for freedom were adjusted using Greenhouse–Geisser epsilon as the conservative correction cue. Interestingly, participants appeared to make Friend decisions much more quickly in the Northern Persian Gulf setting than in the Northern Pacific Ocean. However, no significant main effect for AOR, or for the interaction of AOR and Actual ID, were found.
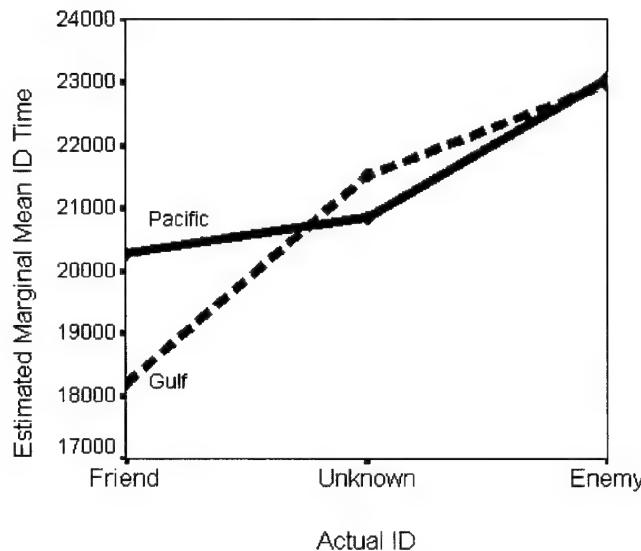
Figure 6. Mean ID response time (ms) for each AOR.

5.2.5.6 Hypothesis 3: Threat Level Response. Figure 7 shows Threat Level plots. All tracks were rated as less threatening in the Northern Pacific Ocean than in the Northern Persian Gulf AOR ($F (1, 16) = 7.932, p = 0.012, \eta^2 = .331$). Friend tracks appeared to be rated as less threatening in the Northern Pacific Ocean as compared to Northern Persian Gulf AOR, but the expected main effect for Actual ID, or for the interaction between AOR and Actual ID, was not significant.
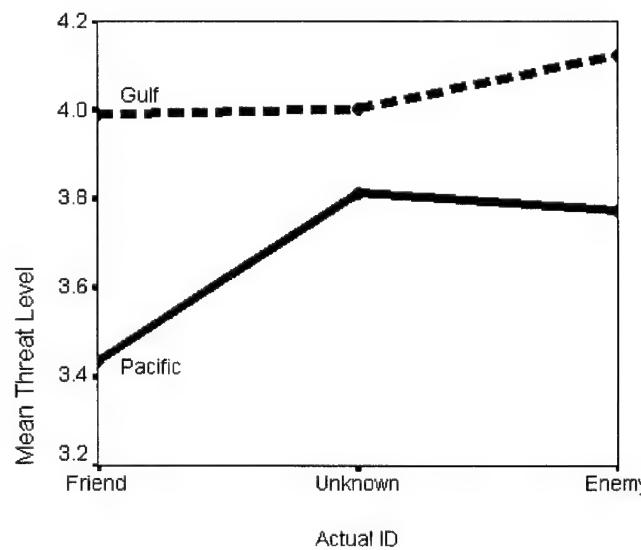
Figure 7. Mean Threat Levels assigned by participants.

23

5.2.5.7 Hypothesis 3: Threat Response Time. Figure 8 shows the results for Threat selection Time. Although Actual ID did not influence Threat Level responses, participants set Threat Levels more quickly when the data were consistent with Friend tracks than with Unknown or Enemy tracks (F (1.367, 21.872) = 8.366, p = 0.005, η^2 = 0.343). This finding indicates that participants were sensitive to ID category, thus lending partial support to Hypothesis 3. The degrees of freedom were adjusted with Greenhouse–Geisser epsilon because the sphericity assumption was violated (Mauchly's W = 0.537, p = 0.009). The main effect for AOR and the interaction between AOR and Actual ID were not significant.
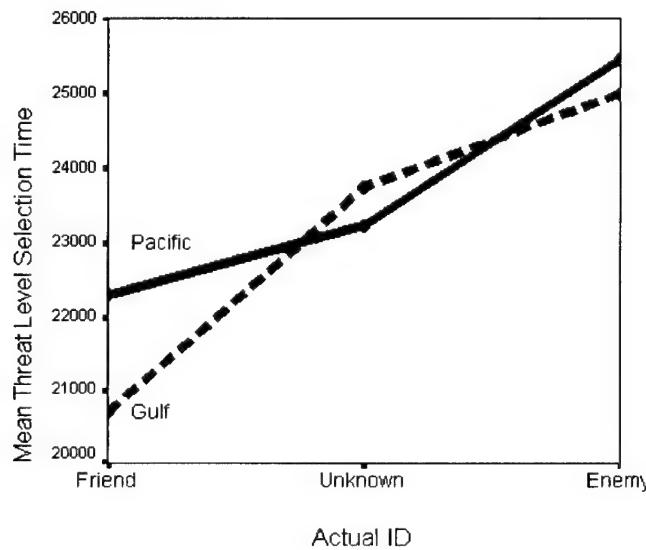
Figure 8. Mean Threat Level selection time.

3.2.6 Study 2 Summary

The results of this study were consistent with our earlier work[9]. The patterns of cue use supported Hypothesis 1 by indicating that participants used fewer cues to evaluate Friendly tracks compared to Enemy tracks. They also used fewer cues to evaluate tracks in the Northern Pacific Ocean than in the Northern Persian Gulf AOR. They did not use all cues equally often. Participants consistently used only 6 to 13 cues out of the 18 available. Of the 6 to 13, six critical cues were identified based on their frequency and sequence of selection by participants. They were Origin, IFF Mode, Intelligence, Altitude, Airlane, and ES.

Participants accurately identified (as Friend, Unknown, or Enemy) a significant majority of the tracks (approx. 70%), thus supporting Hypothesis 2. Accuracy was similar in both AORs and across all ID Categories. Although participants were fairly accurate across all ID categories, they took more time with Enemy tracks than with Friend tracks in both AORs. Tracks were perceived as significantly less threatening in the Northern Pacific Ocean than in the Northern Persian Gulf environment, partly supporting Hypothesis 3, but no differences in Threat Level ratings among Friend, Unknown, or Enemy tracks were found. However, participants did set Threat Levels significantly more quickly when the data were consistent with Friend tracks than with Unknown or

---

[9] Liebhaber and Smith, *ibid.*

Enemy tracks. This indicated that ID category might affect the cognitive processing of the tracks, but not the final decision.

3.3. STUDY 3: EFFECT OF CONFLICTING INFORMATION ON ID AND THREAT DECISIONS

The goal of Study 3 was to empirically evaluate the effect of conflicting information on track assessment. Study 1 provided us with a list of cues and their corresponding influence on threat level. Study 2 demonstrated that participants weigh cues unequally. For example, rather than evaluate all the data, participants tended to rely heavily on a few cues, namely, Origin, IFF Mode, Intelligence, Altitude, Airlane, and ES. If the concepts of cue-threat relationships and cue weights are valid, then interfering with track data should cause a change in threat ratings and identification, and the change should be greater for higher weighted cues than for lower weighted cues. Interference was provided by giving the participant data that conflicted with the ID of the track being evaluated.

3.3.1 Participants

Eighteen U.S. Navy personnel with at-sea AD experience participated in this study. The average at-sea experience across all participants was 2.4 years. Most participants had experience in more than one AD role. Appendix A defines the roles.

3.3.2 Design

Study 3 used a 3 (ID-Platform) X 8 (Conflict Cue) X 3 (Conflict Platform) within subject design.

3.3.2.1 Independent Variables

- **ID-Platform:** The ID and Platform that were given to the participant corresponded to one of the following three ID and Platform combinations:

  1. Neutral COMAIR

  2. Neutral Surveillance (SURVAIR)/Maritime Patrol (MARPAT)

  3. Suspect Tactical/Military (TACAIR)

- **Conflict Cue:** On some trials, participants saw data that conflicted with the ID-Platform (i.e., the cue value was not typical for the given track). This variable indicates the conflicting cue(s). Only the six most important cues could contain conflicting data. Importance was based on the relative weight (defined below) of each cue. Data for cues 7 through 18 were always consistent with the initial ID-Platform of the track. There were eight levels of conflicting data:

  1. No Conflict. Data for all cues are consistent with track ID-PLATFORM.

  2. Low Conflict. Data for cue 6 (ES) conflicts with ID-PLATFORM.

  3. Medium Conflict. Data for cues 3 through 5 (Airlane–Altitude–Airlane) conflict with ID-PLATFORM.

  4. High Conflict. Data for cues 1 through 2 (Origin–IFF Mode) conflict with ID-PLATFORM.

5. Medium+Low Conflict. Data for cues 3 through 6 (Airlane–Altitude–Airlane/ES) conflict with ID-PLATFORM.

6. High+Low Conflict. Data for cues 1 through 2 and 6 (Origin–IFF Mode/ES) conflict with ID-PLATFORM.

7. High+Medium Conflict. Data for cues 1 through 5 (Origin–IFF Mode/Airlane–Altitude–Airlane) conflict with ID-PLATFORM.

8. All Conflict. Data for cues 1 through 6 (Origin–IFF Mode/Airlane–Altitude–Airlane/ES) conflict with ID-PLATFORM.

- **Conflict Platform:** On trials where the data conflicted with the ID-Platform (i.e., Conflict Cue levels 2 through 8), this variable indicated the type of conflicting data according to the following:

if ID-Platform was	then Conflict Platform was
COMAIR	SURVAIR or TACAIR
SURVAIR	COMAIR or TACAIR
TACAIR	COMAIR or SURVAIR

For example, on trials where the ID-Platform was COMAIR, the conflicting data were taken from either SURVAIR or TACAIR (equally balanced across trials).

3.3.2.2 Dependent variables. The variables were as follows:

- ID: Participant's classification of the track.

- ID Match: Proportion of each participant's IDs that matched the Initial ID (from ID-Platform).

- Platform Type: Type of track participant assigned.

- Platform Match: Proportion of each participant's Platform selections that matched the Initial Platform (from ID-Platform).

- Conflict Match: Proportion of each participant's Platform selections that matched the Conflict Platform.

- Threat Level: Degree of threat posed by the track to own ship/battle group as perceived by the participant from the given data. Scale from 1 (No Threat) to 7 (Capable of destroying one or more ships).

- Inferred Intent: Track's most likely Course of Action (COAL) as determined by the participant. Scale from 1 (Friendly) to 7 (Hostile).

- Confidence: Participant's degree of confidence with current assessment. Scale from 1 (0%) to 7 (100%).

- Time: ID, Platform, Threat Level, Intent Assignment, and Total Trial Time.

26

3.3.2.3 Demographic Variables. Background data collected from each participant were as follows:

- Experience (years at sea in an AD job)

- AD role/position

3.3.3 HYPOTHESES

Given that conflicting data should cause interference with the assessment task, the first step in forming a hypothesis was to determine the nature of the conflict. If AD personnel rely on (i.e., give more weight to) certain cues, then interfering with heavily used cues should cause a decline in performance relative to interfering with less used cues. To test this assumption, it was necessary to derive a hypothesized pattern of interference for the conflicting data on performance. Two types of patterns were possible: (1) a nonlinear pattern based on the relative weight of each cue (computed in Study 2), and (2) a linear pattern based on unweighted cues.

3.3.3.1 Nonlinear Interference Pattern. Table 10 shows the expected pattern of interference based on weighted cues. This pattern is based on two findings in Study 2:

1. Participants relied on six critical cues (Origin, IFF Mode, Intelligence, Altitude, Airlane, and ES).

2. Participants appeared to consider the six critical cues in three combinations.

The three combinations were labeled the High (Origin and IFF Mode), Medium (Intelligence, Altitude, and Airlane), and Low (ES only) weight groups. The six cues were assigned a number that corresponded to their weight group (3 = High, 2 = Medium, 1 = Low). A zero indicated conflicting data. These are the values that appear in the High, Medium, and Low columns in Table 10. For example, in the Conflict Cue = 4 condition, the first two cues are represented with zeros. In that condition, those cues (Origin and IFF Mode) conflict with the ID-Platform. The three Medium weight cues receive their full weight (2), as does the final cue (1). The weights were summed to produce a value that represented the quantity of no conflicting data that were available to the participant on a given trial. The sum provided the basis for hypothesized performance; a higher value meant better performance.

Table 10. Calculation of expected interference and hypothesized performance, assuming cues are weighted.

Conflict Cue	Origin & IFF Mode (High)	Intel, Altitude, & Airlane (Medium)	ES (Low)	Sum	Expected Interference	Hypothesized Performance
1 None	3 3	2 2 2	1	13	None	Best
2 Low	3 3	2 2 2	0	12	Low	Good
3 Medium	3 3	0 0 0	1	7	Intermediate	Intermediate
4 High	0 0	2 2 2	1	7	Intermediate	Intermediate
5 Medium+Low	3 3	0 0 0	0	6	Intermediate	Intermediate
6 High+Low	0 0	2 2 2	0	6	Intermediate	Intermediate
7 High+Medium	0 0	0 0 0	1	1	High	Poor
8 All	0 0	0 0 0	0	0	Maximum	Worst

3.3.3.2 Linear Interference Pattern. Table 8 shows the pattern of interference based on unweighted cues. This pattern would be expected if participants relied on each cue equally. All six critical cues were assigned a value of 1, indicating their equal weight. Again, zero indicates conflicting data. The weights were summed for each condition to produce the quantity of nonconflicting data available to the participant on a given trial. The sum weights provided the basis for hypothesized performance; a higher value indicated better performance.

Table 11. Calculation of expected interference and hypothesized performance, assuming cues are not weighted.

Conflict Cue	Origin & IFF Mode (High)	Intel, Altitude, & Airlane (Medium)	ES (Low)	Sum	Expected Interference	Hypothesized Performance
1 None	1 1	1 1 1	1	6	none	best
2 Low	1 1	1 1 1	0	5		
4 High	0 0	1 1 1	1	4	*gradual*	*gradual*
3 Medium	1 1	0 0 0	1	3	*increase*	*decline*
6 High+Low	0 0	1 1 1	0	3	*in*	*in*
5 Medium+Low	1 1	0 0 0	0	2	*interference*	*performance*
7 High+Medium	0 0	0 0 0	1	1		
8 All	0	0	0	0	maximum	worst

28

Using Tables 10 and 11, several predictions about the role of conflicting data can be made. These predictions are described in hypotheses 1 through 3.

3.3.3.3 Hypothesis 1: Overall Role of Conflicting Data. Analysis of Hypothesis 1 will determine if conflicting data lead to changes in ID and threat assessments. When data conflict with the given ID-Platform, participants' responses should move away from the response suggested by the given data (in ID-Platform) and toward the response suggested by the conflicting data (in Conflict Platform). Specifically, conditions of moderate to high conflict (Conflict Cue = 3 through 8) we expect the following:

1. ID to change from Neutral or Suspect to Unknown.

2. Platform to change from COMAIR or TACAIR to MARPAT.

3. Threat Level to change from an extreme to a mid-range value (i.e., high threat levels will decrease and low threat levels will increase).

4. Inferred Intent to change from an extreme to a mid-range value (Uncertain Intent).

5. Confidence level to change from an extreme to a mid-range value (Unsure Confidence).

Support for Hypothesis 1 will be indicated by a significant ID-Platform by Conflict Cue interaction. Conflict Cue indicates the role of a particular type of conflicting information. Its interaction with ID-Platform is expected because for any given dependent variable, different responses are anticipated for different ID-Platform combinations. For example, in the presence of no conflicting data (Conflict Cue = 1), a Suspect TACAIR should be rated as more threatening than a Neutral COMAIR. Then, as more conflict is introduced, the Threat Level of the Suspect TACAIR should decrease. (Recall that the conflicting data fit a less-threatening MARPAT or COMAIR profile.) The opposite effect should be seen when participants are given a Neutral COMAIR track to evaluate.

3.3.3.4 Hypothesis 2: Data Quantity versus Weight. Hypothesis 2 will evaluate the (linear or non-linear) relationship among the Conflict Cues. If participants were relying on the quantity of information available rather than weighting information, then we would expect the following:

1. Proportions of Proportion of ID and Platform responses that match the Given ID-Platform should gradually decline as the amount of conflicting data increases (Conflict Cue from 1 to 8).

2. Proportion of Platform responses that match Conflict Platform should show a corresponding linear increase as the amount of conflicting data increases.

3.3.3.5 Hypothesis 3: Specific Patterns of Interference. There is little guidance to define the exact nature of the interference other than the cue weights from Study 2. However, the question remains, Which cues cause the most interference? Based on cue weights, conflicting data in the highly weighted cues should cause the most interference. Therefore, we expect significant differences between low weighted (Conflict Cue = 1 and 2) conditions) and higher weighted (Conflict Cue = 3 to 8) for proportions of ID, Platform, and Conflict Matches.

3.3.4 Procedure

Participants filled out an on-screen questionnaire about their experience and participated in a computer-based experiment. The computer displayed information about an air track. On each trial, participants were given a track identification (ID-Platform) and all 18 cues. Their task was to confirm the ID and Platform shown on the screen and then assign a level of threat to the track. Participants evaluated one track at a time. They were given three practice trials and then 45 experimental tracks were random. The research session lasted about 30 to 45 minutes per participant.

3.3.4.1 Display. Figure 9 shows a sample screen from one trial. From top to bottom, the display components were as follows:

- Track ID: The Given-ID (Suspect TACAIR, Suspect MARPAT, or Neutral COMAIR).

- Trial Number: Current trial number.

- Cues: A 3 X 6 grid of data that were always in view during the trial. The cues were always in the same positions.

- ID through Confidence: Participants responded by clicking on one of the buttons for each item. Once selected, they could not change the previous response.

- Next: Participants clicked this button to go to the next trial.

Figure 9. Screen from one trial for Study 3.

3.3.4.2 Sequence of Events

1. Participants were briefed about the task and given the following instructions to read. The instructions were available to the participant during the experiment.

   a. **Instructions**

   Situation:

   - You are in the Northern Persian Gulf.

   - You are the Anti-Air Warfare Coordinator (AAWC) onboard an Aegis-class cruiser.

   - You have just come on duty and must assume the watch.

   - You have not received a "Pass Down" or other briefing.

   - You have a series of air tracks to evaluate.

   ID Tasking:

   - An initial Track ID has been suggested.

   - Available supporting information is displayed.

   - Country of Origin indicates point of initial detection.

   ADW Mission and Tasking:

   - Maintain a Unit Self-Defensive AAW Posture.

   - Conduct Positive Identification (PID) by evaluating and correlating available data.

   - Assess the level of threat posed by each track to own-ship.

   ID Criteria for AIR:

   Note: SURVAIR=MARPAT

CHARACTERISTICS/BEHAVIORS	PRI ID:
A. Any GENERAL ACFT/HELO meeting non-hostile profile, or meeting COMAIR profile.	NEUTRAL
B. Any track with a profile that does not presently permit ID.	UNKNOWN
C. Any MIL/TAC/SURV AIR or GENERAL ACFT/HELO meeting potentially hostile profile.	SUSPECT

2. Participants were then shown the components of the display and were told:

   a. Look at the Initial ID for each new experimental trial.

   b. Select an ID, Platform, Threat Level, Intent, and Confidence Level.

31

     i.   Intent is your perception of the track's intentions.

     ii.   Confidence is your confidence in the above responses.

  c. Once you make a selection, you cannot return to it.

  d. After you select a Threat Level, the computer will automatically display the next track.

3. Please work as quickly and as accurately as possible.

4. Tracks 1, 2, and 3 are practice trials.

5. Participants went through three practice trials. After the practice trials, participants' remaining questions were answered.

6. When the participants were ready, the experimental trials began.

3.3.4.3 Trial Input Data File. The input data file contained the information that was displayed on the screen for each trial. The file contained three practice trials and 45 experimental trials. The experimental trials were random. The track data for each cue was constructed as consistent or inconsistent with the Initial ID. Data values for each of the 18 cues were chosen to create plausible track behaviors for each trial. Slight modifications to the data (e.g., 15,000' versus 15,500') were used to create multiple trials for the same type of track (e.g., Neutral COMAIR).

3.3.5 Results

3.3.5.1 Hypothesis 1. A 3 (ID-Platform) X 8 (Conflict Cue) X 3 (Conflict Platform) within subjects Multivariate ANOVA was conducted. The results are reported by dependent variable: ID and ID Time, Platform Type and Platform Time, Threat Level and Threat Time, Inferred Intent and Intent Time, and Confidence Level and Confidence Time.

3.3.5.1.1 ID Assignment and Time. Appendix E shows mean ID assignments and ID times. Hypothesis 1 was supported for ID assignment. As expected, participants' ID assignments matched the given ID-Platform when there was relatively little conflicting data (None, Low, and Medium conditions). In conditions with moderate to high levels of data conflict, participants were more likely to assign Unknown IDs to track. The interaction between ID-Platform and Conflict Cue was significant (*Wilks' Lambda* (12,3) = 0.005, p = 0.004), indicating that Suspect tracks moved down towards Unknown, and Neutral tracks moved up toward Unknown. Figure 10 shows this relationship. There were also significant main effects for ID-Platform (*Wilks' Lambda* (2,13) = 0.378, p = 0.002) and for Conflict Platform (*Wilks' Lambda* (1,14) = 0.365, p = 0.000). These effects indicated that participants were sensitive to changes in the source of the data that they were evaluating (i.e., data were from either the given or conflicting platform). ID Time did increase with increasing levels of conflicting data, but the effect was not significant.
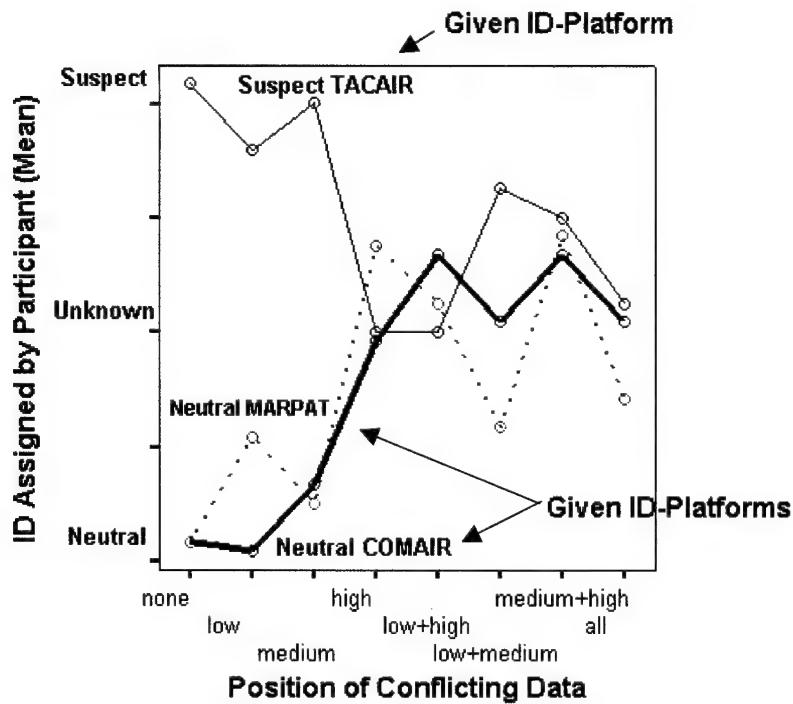
Figure 10. Participant's ID assignments for given positions of conflicting data (lines indicate the type of given ID-Platform).

3.3.5.1.2 Platform Assignment and Time. Appendix E shows Mean Platform assignments and times. As expected, participants' Platform Assignments matched the Given ID-Platform when there was relatively little conflicting data, and they changed their Platform Assignments when there were conflicting data (*Wilks' Lambda* (12,3) = 0.019, *p* = 0.030). Figure 11 shows this relationship. For example, when the given ID-Platform was COMAIR, participants were more likely to call the track a HELO or MARPAT in the presence of conflicting data.

Not only did participants change Platform Assignments, their new Platform Assignment matched that of the Conflict Platform. The ID-Platform and Conflict Platform interaction (*Wilks' Lambda* (2,13) = 0.605, *p* = 0.038) is in Figure 12. For example, the left panel of Figure 12 shows that when the given Platform was COMAIR and the conflicting data were from a MARPAT track, participants tended to respond HELO. When the data were from a TACAIR, participants responded HELO or MARPAT.

Figure 13 shows Mean Platform decision times. Participants took significantly longer to select a platform (COMAIR, HELO, MARPAT or TACAIR) when there were conflicting data than when there were no conflicting data (*Wilks' Lambda* (6,9) = 0.260, *p* = 0.026).

33
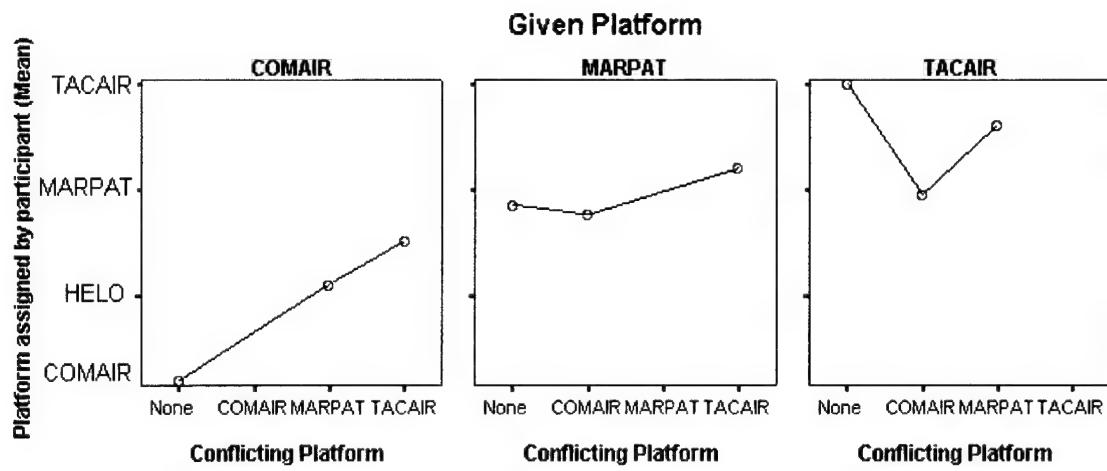
Figure 11. Participant's Platform assignments for given positions of conflicting data (lines indicate the type of given ID-Platform).
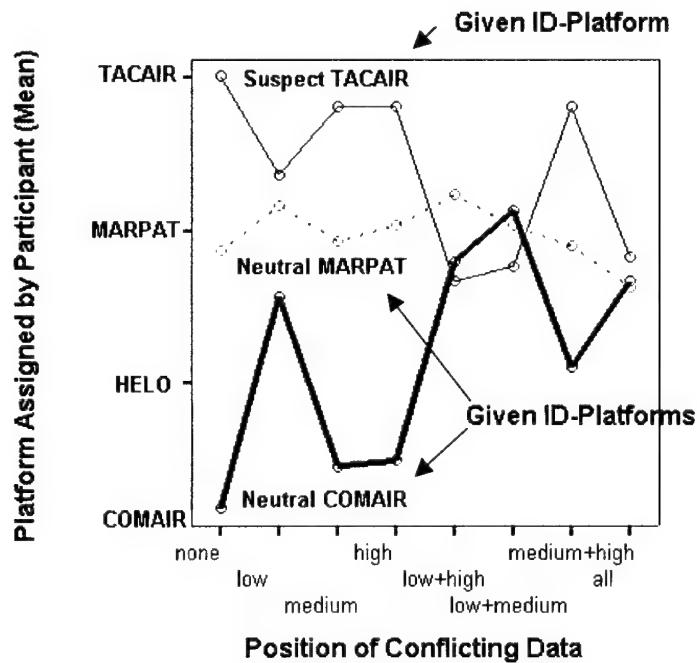
Figure 12. Participant's Platform assignments for given type of conflicting data (COMAIR, MARPAT, TACAIR) for each type of given ID-Platform).

34
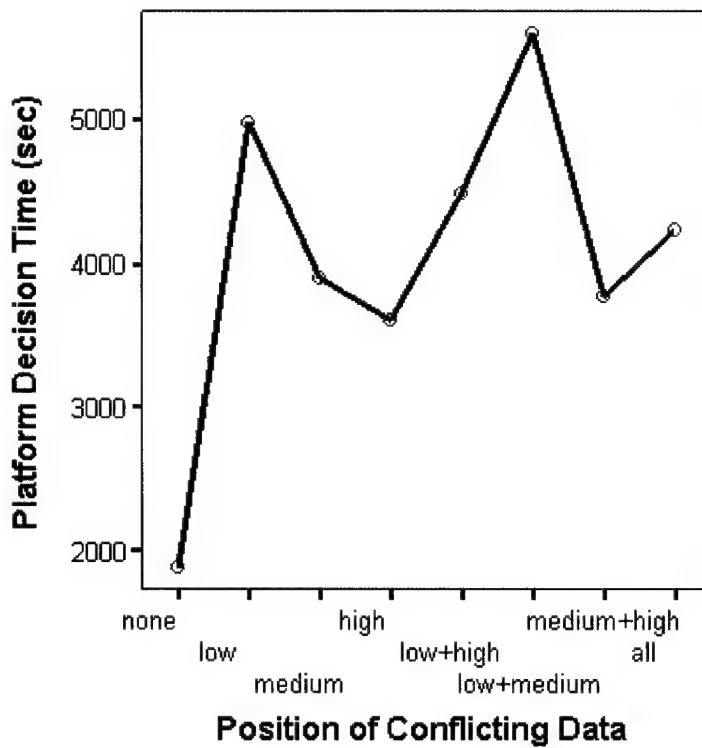
Figure 13. Participant's Platform assignment time for each position of conflicting data.

3.3.5.1.3 Threat Level Assignment and Time. The interaction between Conflict Cue and given ID-Platform was not significant, which indicated that Hypothesis 1 was not supported for Threat Level assignment. Appendix E and Figure 14 show the means. There were significant main effects for ID-Platform (*Wilks' Lambda* $(2,13) = 0.249$, $p = 0.000$) and for Conflict Platform (*Wilks' Lambda* $(1,14)$ $= 0.241$, $p = 0.000$). The findings for Threat Selection Time did not support Hypothesis 1. Appendix E shows the means.

Further analysis indicated that Threat Level was sensitive to Conflict Cue. Paired t-tests were conducted using Threat Level as the dependent measure and Conflict Cue (Low versus All) as the grouping factor. Separate analyses were run for each type of Conflicting Platform (COMAIR, MARPAT, and TACAIR), and are plotted in Figure 15. Significance level was altered using the Boneferroni correction to control test wise error rate: alpha $= 0.05/3 = 0.017$. When the given ID-Platform was Neutral COMAIR, Threat Levels were significantly lower in the Low conflict condition compared to the All conflict condition (t $(14) = 2.82$, $p = 0.014$). Threat Levels increased when the data became less like COMAIR and more like MARPAT and TACAIR. The same was true for TACAIR. Threat Levels decreased as the data looked more like MARPAT or COMAIR (t $(14) =$ -4.14, $p = 0.001$). MARPAT Threat Levels also increased slightly in the All data conflict condition compared to the Low data conflict condition (MARPAT: t $(14) = -3.23$, $p = 0.006$). Figure 16 clearly shows that Hypothesis 1 is supported for COMAIR and TACAIR. The Threat Level in those two situations moves toward a middle threat value.
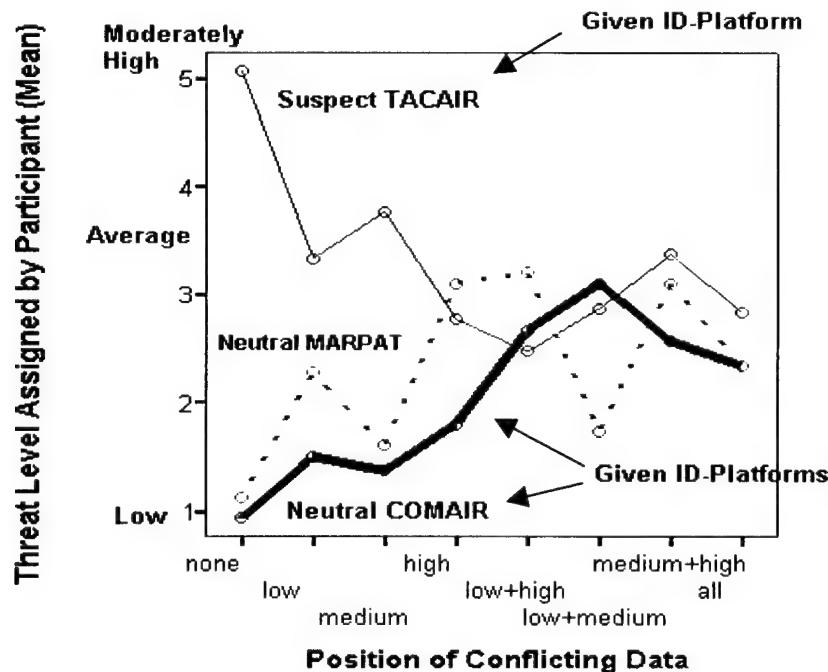
Figure 14. Participant's Threat Level assignments for given positions
of conflicting data (lines represent given ID-Platform).
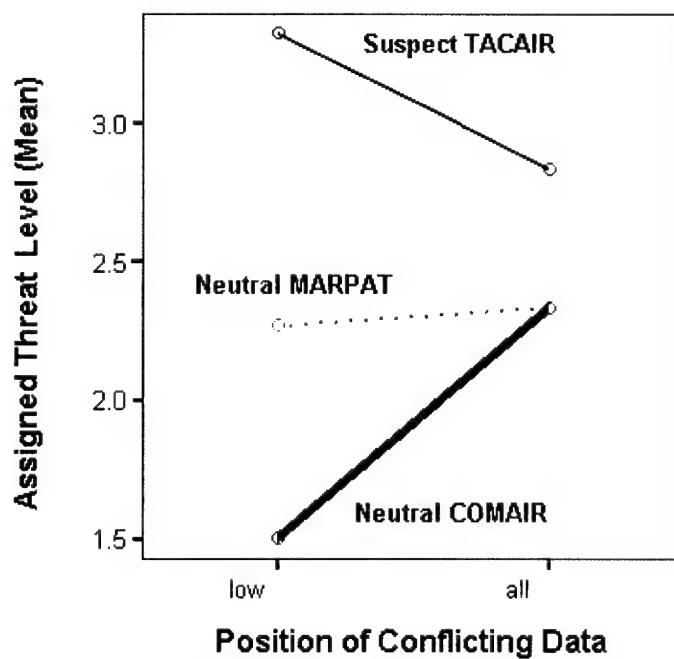
Figure 15. Participant's Threat Level assignments for Low and High (All)
levels of conflicting data (lines indicate each type of ID-Platform).

3.3.5.1.4 Intent Assignment and Time. There was a significant interaction between ID-Platform and Conflict Cue (*Wilks' Lambda* (12,3) = 0.004, *p* = 0.003), thus, supporting Hypothesis 1. Tracks of all types were more likely to be assigned a midscale value (e.g., Uncertain Intent) as the amount of conflicting data increased. Appendix E and Figure 16 show this relationship. The interaction between ID-Platform and Conflict Platform was also significant (*Wilks' Lambda* (2,13) = .575, *p* = 0.027), indicating that participants were sensitive to the conflicting data. Participants assigned higher (i.e., more hostile) intent levels to tracks when the conflicting data came from a potentially threatening track (e.g., TACAIR).

For Intent selection time, the ID-Platform by Conflict Platform interaction was not significant. The Conflict Platform (*Wilks' Lambda* (1,14) = 0.751, *p* = 0.049) main effect was significant. On average, participants took the longest when there were no conflicting data (None condition). Interestingly, almost all the Intent Time delay in the None condition occurs when the participants are told that the track is a Suspect TACAIR (see Figure 17).
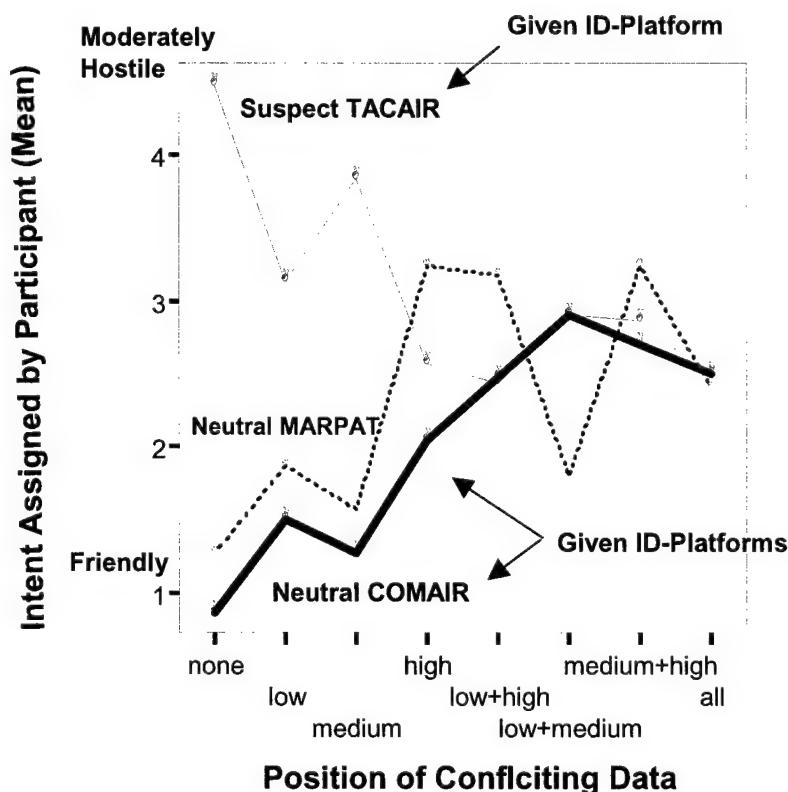
Figure 16. Participant's Intent assignments for given positions of conflicting data for each type of given ID-Platform.
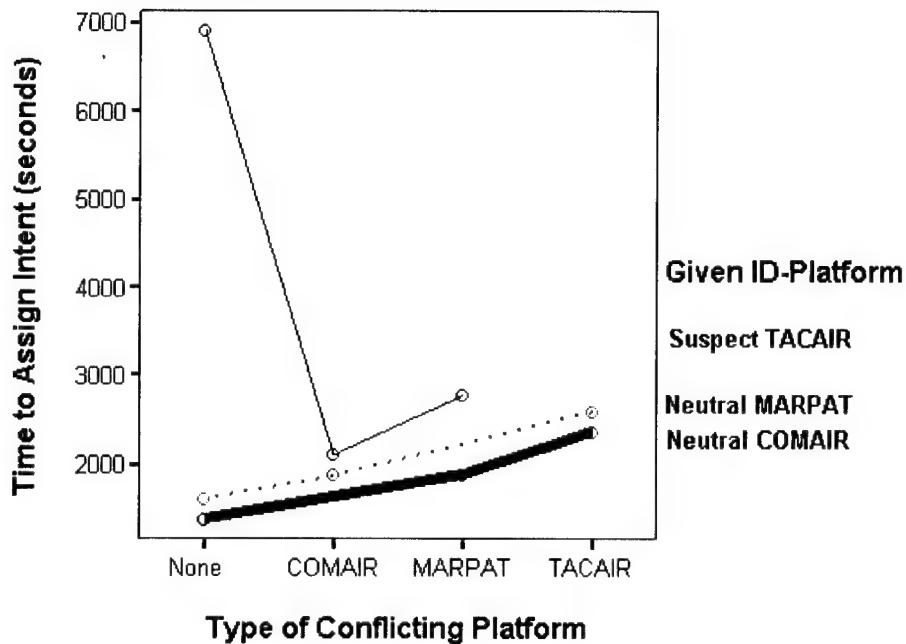
Figure 17. Intent Time for combinations of given ID-Platform and Platform of conflicting data (lines represent given ID-Platform).

3.3.5.1.5 Confidence Level Assignment and Time. Appendix E shows Mean Confidence Level assignments and times. There were no significant effects for Confidence Level assignment or selection time. Confidence was highest when there were no conflicting data and lowest in the Medium+High Conflict Cue condition. Figure 18 shows this relationship.
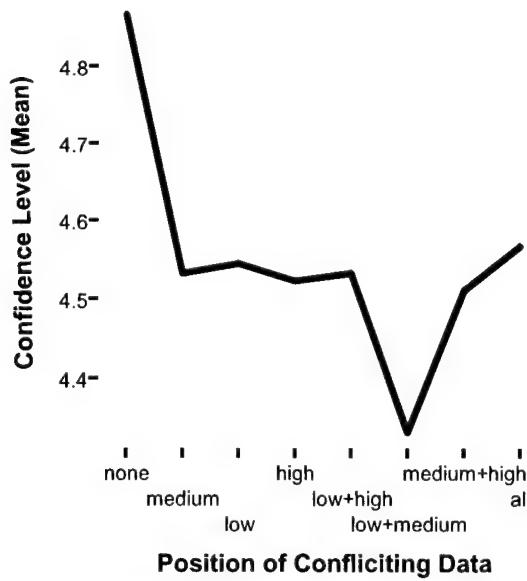
Figure 18. Relationship between participant's confidence level and conflicting data.

3.3.5.2 Hypothesis 2. The linear component of a polynomial contrast from three univariate ANOVAs was used to evaluate Hypothesis 2. The dependent variables were ID Match, Platform Match, and Conflict Match. The independent variable was Conflict Cue. Significance level was altered using the Boneferroni correction to control test wise error rate: alpha = 0.05/3 = 0.017. Table 12 shows the proportion of responses for each dependent variable. The proportions indicate a participant's likelihood to stay with the Given ID-Platform or to switch to another Platform. As expected, the proportion's ID matches and Platform matches did decline from low to high amounts of conflicting data, but not in a linear manner (see Figure 19). The lack of support for Hypothesis 2a indicated that each participant's ID and Platform decisions were related to conflict with specific cues, and not just the quantity.

Table 12. Overall likelihood to stay with or switch from the given ID-Platform type.

Position of Conflicting Data	Likelihood to Stay with		Likelihood to Switch to
	Given ID (ID Match = Yes) (%)	Given Platform (Platform Match = Yes) (%)	Conflicting Platform (Conflict Match = Yes) (%)
No Conflicting Data	82.3	95.6	0.00
Low (ES)	85.0	49.3	42.2
Medium (Airlane–Altitude–Airlane)	88.7	85.3	10.0
High (Origin–IFF Mode)	42.0	84.4	4.4
Low+High (Origin–IFF Mode/ES)	35.6	46.7	44.4
Low+Medium (Airlane–Altitude–Airlane/ES)	67.1	34.4	52.2
Medium+High (Origin–IFF Mode/Airlane-Altitude–Airlane)	41.3	73.3	11.1
All (Origin–IFF Mode/Airlane–Altitude–Airlane/ES)	54.9	44.7	48.9

Participants seemed especially sensitive to conflicting Origin, IFF Mode, and ES. When making an ID (i.e., Neutral, Suspect, Unknown) decision, participants were most likely to switch ID when Origin and IFF Mode conflicted with the Given ID-Platform. The left panel of Figure 19 shows that effect. The lowest points on the curve are Conflict Cues that contain Origin and IFF Mode (i.e., High, Low+High, Medium+High, & All). When participants made a Platform (i.e., TACAIR, COMAIR, MARPAT, and HELO) decision, they were most likely to switch Platforms when ES conflicted with the Given ID-Platform. The right panel of Figure 19 shows the ES effect. The lowest points on the curve are Conflict Cues that contain ES (i.e., Low, Low+High, Low+Medium, and All).

The number of Conflict Matches increased, also as expected, but the linear component was significant (*Linear estimate* = -0.174, $p = 0.004$), thus supporting Hypothesis 2b. Figure 20 shows the relationship. The curve is the inverse of the Platform curve in the right panel of Figure 19. The highest points in Figure 20 indicated a high likelihood to switch Platforms, and correspond to conflicting ES data. Recall that conflicting data for the Given ID-Platform was valid data from another type of Platform. A Conflict Match indicated that participants not only switched from the Given ID-Platform, but that the Platform that they choose matched the Platform that the conflicting data were from. It appeared that although participants were evaluating specific cues, the quantity of conflicting information eventually led them to switch from one Platform to another.
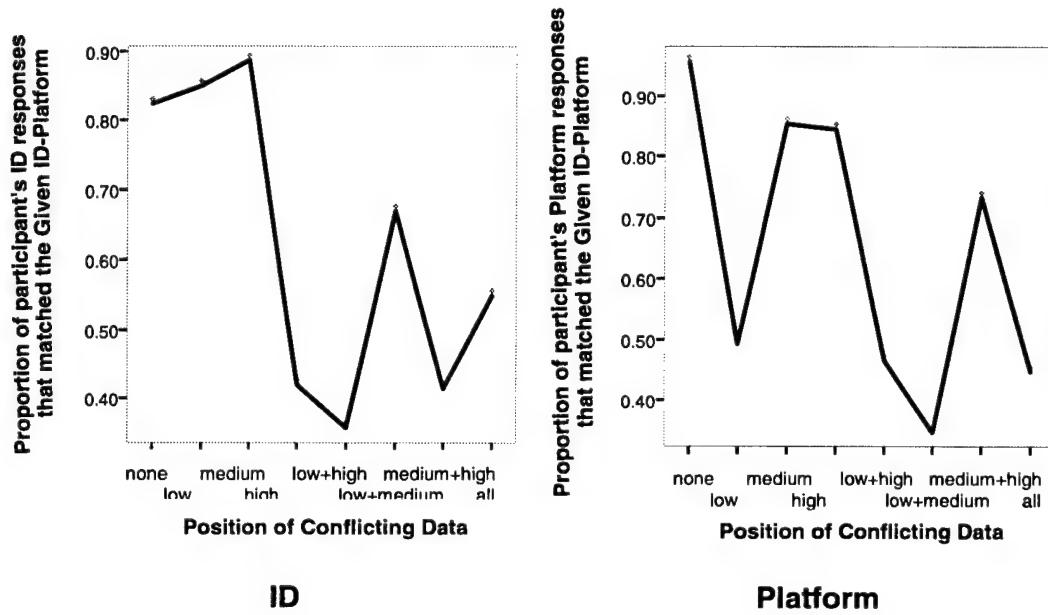
Figure 19. Proportion of ID (left graph) and Platform (right graph) responses that matched the Given ID-Platform at each level of conflicting data.

40
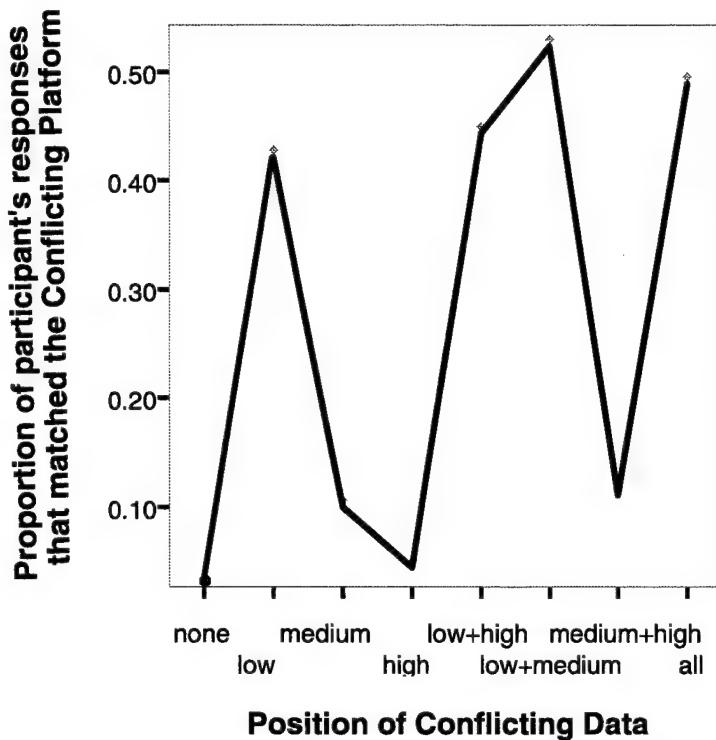
Figure 20. Proportion of Platform responses that matched the Conflicting Platform at each level of conflicting data.

3.3.5.3 Hypothesis 3. ID, Platform, and Conflict Match proportions were analyzed to determine which cues (as defined by Conflict Cue) led to the most interference. For Hypothesis 3a, the No Conflict condition was compared to all other conditions with the McNemar ordered pairs test. The pattern of significance in Table 10 indicates that all but minimal amounts of conflicting data interfered with ID and Platform assignment. The ID results were as expected. ID decisions appeared to be most affected by high weighted conflicting data: Origin, IFF Mode, Intelligence, Altitude, and Airlane. However, the results for Platform were not as clear. Platform decisions were most affected by the relatively low weighted ES cue. The remaining cues affected decision-making only in combination with ES or with each other.

Table 13. Proportion of trials where the participant's ID and Platform matched the given ID.

Position of Conflict	Proportion of ID Matches (%)	McNemar Significance	Proportion of Platform Matches (%)	McNemar Significance
None	82.3		95.6	
Low	85.0	0.453	49.3	0.007
Medium	88.7	0.219	85.3	0.109
High	42.0	0.000	84.4	0.453
Low+high	35.6	0.013	46.7	0.000
Low+medium	67.1	0.063	34.4	0.000
Medium+high	41.3	0.001	73.3	0.012
All	54.9	0.001	44.7	0.000

The Low condition was compared to all other conditions with the McNemar ordered pairs test for Hypothesis 3b. Table 14 shows the results. Low was used as the basis for comparison because the conflicting platform was never selected when there were no conflicting data (None condition). Conflicting data in the Low position (ES) appeared to cause the most interference. The highest proportion of switching to the conflicting platform occurred in every condition where it was present: Low, Low+High, Low+Medium, and All. There was just under a 50-50 chance that a participant would switch platforms based on data from this one cue.

Table 14. Proportion of trials where the participant's Platform assignment matched (Yes) or did not match (No) the Conflicting Platform.

	Match Conflict Platform			
Position of Conflict	Yes	No	Proportion Yes (%)	McNemar Significance
None	0	45	0.00	
Low	38	52	42.2	
Medium	9	81	10.0	0.000
High	4	86	4.4	0.000
Low+medium	47	43	44.4	0.188
Low+high	40	50	52.2	0.860
Medium+high	10	80	11.1	0.000
All	44	46	48.9	0.391

3.3.6 Study 3 Summary

The results indicated that conflicting data interfered with threat assessment. When there were relatively little conflicting data (None, Low, and Medium Conflict conditions), participant's ID and Platform assignments matched the combination in the Given ID-Platform. However, when participants were in conditions with moderate to high levels of data conflict, they were more likely to change ID, Platform, and Threat Level. Not only did participants change Platform Assignments, when they switched, the new Platform Assignment matched that of the Conflict Platform. Participants took significantly longer to select a platform when there were conflicting data than when there were no conflicting data.

Conflicting data also affected Threat Level and Intent ratings. For COMAIR tracks, Threat Levels increased when the data became less like COMAIR and more like MARPAT or TACAIR. TACAIR Threat Levels also decreased as the data became more like MARPAT or COMAIR. Participants became more uncertain about the intent of a track as the amount of conflicting data increased.

Participant's confidence in their decisions was highest when there were no conflicting data and lowest in the Medium+High Conflict Cue (i.e., Origin+IFF Mode/Intelligence+Altitude+Airlane), but the effects were not significant. On average, participants took longer to respond to all dependent variables for Hypothesis 1 when there were conflicting data. However, response time was not a major factor overall.

The results from Hypotheses 2 indicated that when participants made ID and Platform assignments, they were sensitive to the type of conflicting data and not simply the quantity of information. Participants were most likely to switch ID when Origin and IFF Mode conflicted with the Given ID-Platform. They were likely to switch Platforms when ES conflicted with the Given ID-Platform.

Comparison of the conflicting data condition to the No Conflict condition in Hypothesis 3 indicated that all but minimal amounts of conflicting data interfered with ID and Platform assignment. Participants were most likely to switch track ID (i.e., Neutral, Suspect, Unknown) when there were conflicting, high weighted data: Origin, IFF Mode, Intelligence, Altitude, and Airlane. Platform decisions were most affected by conflicting ES data. Participants were significantly more likely to switch platforms when data from just this one cue conflicted with the given ID-Platform.

43

4. DISCUSSION

These studies were conducted to better understand the relationship between the data available to the AD team and the level of threat they perceive for different types of aircraft in different operational settings. The experienced AD participants were clearly integrating information that, in a shipboard setting, would come from many different sources. Study 1 indicated that participants had developed task knowledge of types of threats, weapons, and the effects of situational context (e.g., ship's mission and regional geopolitical climate) and U.S. Navy doctrine (e.g., Rules of Engagement and Operational Task Orders). Cue weights and patterns of interference from studies 2 and 3 suggested that participants relied on heuristics to assess aircraft.

The first study provided default threat levels for aircraft and weapon systems. The study also listed the relationship between specific values of cues (factors used by AD personnel to evaluate aircraft) and the corresponding perceived threat ratings (e.g., an altitude of 12,000 ft corresponds to a 0.5 increase in threat level) in littoral and open-ocean waters. Experienced U.S. Navy AD personnel provided the ratings. The results included the cues, range of data values that each cue could assume, and the corresponding increase or decrease in default threat level for each data value. Not surprisingly, aircraft in the Friendly ID category were rated as less threatening than aircraft in the Hostile ID category. Participants rated unknown aircraft tracks as an intermediate threat. Regardless of ID Category, aircraft in the Littoral AOR setting were rated as more threatening than aircraft in the Open-Ocean AOR setting.

Study 2 empirically established the relative importance of each cue. Important cues were defined as cues providing the most evidence or information value to the assessment process. We presumed that cues that were evaluated earlier and more often were more important than later and less frequently selected cues. On average, partiipants used between 8 and 11 cues out of the 18 available, and they selected the cues in roughly the same sequence for each type of track (i.e., Friend, Unknown, or Enemy). Participants used fewer cues to evaluate Friendly tracks compared to Enemy tracks. They also used fewer cues to evaluate tracks in the Open-Ocean than the Littoral AOR. Aircraft in the Open-Ocean AOR were perceived to be significantly less threatening than aircraft in the Littoral AOR.

Participants relied on some cues more than other cues, but selection order or frequency did not adequately capture the notion of cue importance. To quantify the concept of importance, an index of Relative Cue Weight (W) based on selection order and frequency was computed for each cue. The resulting index number represented the relative importance of each cue as a function of how often it was picked in a given order. Of the 18 cues, 6 cues were considered critical based on W score. They were Origin, IFF Mode, Intelligence, Altitude, Airlane, and ES.

The third study empirically evaluated the effect of conflicting information on track ID, Platform assignment, and threat assessment. The findings indicated that participants' decisions were influenced by conflicting data in just one or more critical cues and not simply the overall quantity of conflicting information. Participants were given an aircraft ID-Platform identification (e.g., Neutral MARPAT) and all 18 cues on each trial. Data for the six critical cues identified were either consistent with or conflicted with the given ID-Platform. Conflict was systematically varied across the six cues. Participants viewed the information and then assigned an ID, Platform Type, and Threat Level. The overall affect of conflicting data indicated that all but minimal amounts of conflicting data interfered with ID and Platform assignment and with threat assessment. Participants were more likely to change from the Given ID and Platform in conditions with moderate to high levels of data

45

conflict. If they did change from the given information, the new ID, Platform, and corresponding Threat Level were consistent with those indicated by the conflicting data. If the given platform was COMAIR and the conflicting data were from a TACAIR, participants called the track a TACAIR or MARPAT and assigned a higher Threat Level.

It was not surprising that conflicting data prompted participants to change the ID and Platform. The interesting finding was that conflicting data in specific cues caused interference. Participants were most likely to switch track ID (i.e., Neutral, Suspect, Unknown) when there were conflicting, high-weighted data: Origin and IFF Mode, and to a lesser extent, Intelligence, Altitude, and Airlane. They were likely to switch Platforms (i.e., COMAIR, MARPAT, and TACAIR) when ES conflicted with the Given ID-Platform. Participants were significantly more likely to switch platforms when data from just this one cue conflicted with the given ID.

The findings of this program are consistent with some expectations that grew out of the TADMUS program (Morrison, 2000). Results indicate that experienced team members form a hypothesis (i.e., activate a template that corresponds to a particular type of aircraft), evaluate the evidence (i.e., cues), and then make a plausible assessment based on the (supporting and contradicting) data.

However, the findings appear to be inconsistent with some aspects of explanation-based reasoning and related work that shows that decision-making is based on recognition of similar patterns of behavior or situations (e.g., Klein, 1997; Marshall, Christensen, and McAllister, 1996). Clearly, participants were not evaluating all the available evidence in all situations. We found that participants use different cues and different numbers of cues in different situations and that the data in one or two important cues could cause significant changes in ID and Threat Level. This finding implied that participants were either evaluating individual and small clusters of cues or had an enormous number of memorized patterns and pattern recognition networks.

The current findings are also complementary with previous studies. The inconsistency may be due to the nature of the tasks. Our studies asked participants to specifically identify the cues they were evaluating and how the data would affect their perception of the track. Other research (e.g., Marshall, Christensen, and McAllister, 1996) has correlated a fixed number of cues with course of action decision-making behaviors (e.g., ignore, intercept, warn, etc.). While our research has dealt with the process of gathering and assimilating information into a coherent threat picture, the work of Klein (1997); Marshall (2000); Marshall, Christensen, and McAllister (1996); and others (e.g., Cohen et al., 1993; Endsley, 1995; Zachary et al., 1992) has focused on recognizing the pattern of information in the assimilated threat picture, incorporating goals and planning, and forming an executable course of action. Our findings have detailed the initial steps in the threat awareness process, which roughly correspond to Levels 1 and 2 in Endsley's (1995) Situation Awareness model. The results of the decision-making oriented research correspond to later stages in Endsley's model. Additional work is needed to determine the relationship of the current findings to situation awareness models in the naturalistic decision-making framework.

5. CONCLUSION

Results of these studies support the notion that the threat assessment process involves activating an initial platform schema and then comparing track data to each default value for each cue of that schema. The schemas contain a default threat level for a particular type of aircraft (e.g., TACAIR) and an ordered list of cues to evaluate. The evaluation process increases or decreases the default threat level, depending on the data for each cue (e.g., high altitude will decrease threat level). For some cues, the amount of increase or decrease depends on the location of the ship (i.e., its AOR). Evaluation order is based on the weight or importance of each cue. Heavily weighted cues are evaluated earlier than lower weighted cues.

Threat assessment clearly requires a high level of tactical experience. However, experience related bias might be easily introduced into the process. The experienced AD personnel in our studies did not choose to evaluate all available cues and they relied on some cues more heavily than on other cues. Decision Support Systems and displays need to consider these behaviors and their potential to introduce biases. A complete threat assessment model should evaluate all cues, display the threat level and contribution of each cue to the overall threat level, and provide input to a system that suggests an appropriate course of action. Understanding this process will lead to better guidelines for tactical situation displays. Effective displays will present information consistent with the threat assessment task and the user's mental model of that task.

6. REFERENCES

Bell, H. H. and D. R. Lyon. 2000. "Using Observer Ratings to Assess Situation Awareness." In *Symposium Proceedings, Threats, Countermeasures, and Situational Awareness: Teaming for Survivability Symposium and Exhibition.* CD-ROM. Electronic Warfare Advanced Technology Program. Virginia Beach, VA. June.

Chalmers, B. A. 1998. "On the Design of Computer-Based C2 Decision Support for a Modern Frigate." In *Proceedings of the 4th International Command and Control Research and Technology Symposium,* U.S. Department of Defense, C4ISR Cooperative Research Program. Nasby Park, Sweden. 14–16 September.

Chalmers, B. A., J. R. Easter, and S. S. Potter. 2000. "Decision-Centered Visualizations for Tactical Decision Support on a Modern Frigate." In *2000 Command and Control Research and Technology Symposium.* CD- ROM. Naval Postgraduate School, Monterey, CA.

Cohen, M. S., L. Adelman, M. A. Tolcott, T. A. Bresnick, and F. A. Marvin. 1993. "Cognitive Framework for Battlefield Commander's Situation Assessment." Technical Report 93-1. Cognitive Technologies, Inc., Arlington, VA.

Endsley, M. R. 1995. "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors, 37,* 32–64.

Gaito, J. 1980. "Measurement Scales and Statistics: Resurgence of an Old Misconception," *Psychological Bulletin, 87,* 564–567.

Kelly, R. T. and R. A. Moore. 1996. "*Situational Parameters and Scenario Events Associated with Decision-Making Workload.* TADMUS Technical Note (Dec). Pacific Science & Engineering Group, Inc., San Diego, CA 92122.

Kirk, R. E. 1995. *Experimental Design: Procedures for the Behavioral Sciences.* Brooks/Cole Publishing Company, Pacific Grove, CA.

Klein, G. A. 1993. "A Recognition-Primed Decision (RPD) Model of Rapid Decision Making." In *Decision Making in Action: Models and Methods* (pp. 138–147). G. A. Klein, J. Orasanu, R. Caldewood, and C. Zsanbok, Eds. Ablex Publishing Corp., Norwood, NJ.

Klein, G. A. 1997. "The Recognition-Primed Decision (RPD) Model: Looking Back, Looking Forward." In *Naturalistic Decision Making* (pp. 285–292). G. A. Klein, J. Orasanu, R. Caldewood, and C. Zsanbok, Eds. Lawrence Erlbaum Associates, Inc., Mahwah, NJ.

Marshall, S. 2000. "Decision-Making Schemas in Rapidly Changing Situations." [On-line: http://www.sci.sdsu.edu/cerf/content/tadmus.html]. San Diego State University, Cognitive Ergonomic Research Facility, San Diego, CA 92120. November.

Marshall, S. P., S. E. Christensen, and J. A. McAllister. 1996. "Cognitive Differences in Tactical Decision Making." In *Proceedings of the 1996 Command and Control Research and Technology Symposium* (pp. 122–132). Naval Postgraduate School. Monterey, CA.

Michell, J. 1986. "Measurement Scales and Statistics: A Clash of Paradigms," *Psychological Bulletin, 100,* 398–407.

Morrison, J. G. 2000. *Tactical Decision Making Under Stress (TADMUS) Command 21: Decision Support for Operational Command Centers (DeSOCC)* [On-line: http://www-tadmus.spawar.navy.mil]. SPAWAR Systems Center, San Diego, Code 244, 53560 Hull St., Bldg. A33, Rm. 1405, San Diego, CA 92152-5001. August.

Osga, G. A. 1999. July/August). *The Task-Centric Watchstation. Surface Warfare Magazine* [On-line: http://surfacewarfare.nswc.navy.mil/magazine/ja_toc.html].

Pennington, N. and R. Hastie. 1988. "Explanation-Based Decision Making: Effects of Memory Structure on Judgment," *Journal of Experimental Psychology, 14*, 521–533.

Pennington, N and R. Hastie. 1993. "Reasoning in Explanation-Based Decision Making," *Cognition, 49*, 123–163.

Schulze, K.G., L. B. Achille, A. Schmift-Nielsen, B. Deubner, and J. Stroup. 1999. "A Networked Wizard-of–Oz Experiment to Study the Effect of Communication on Decision-Making Performance." Technical Report. NRL/FR/5510-99-9901. Naval Research Laboratory, Navy Center for Applied Research in Artificial Intelligence, Information Technology Division, SPAWAR, 4301 Pacific Highway, San Diego, CA 92110-3127.

Smith, D.E. and S. P. Marshall. 1997. "Applying Hybrid Models of Cognition to Decision Aids." In C.E Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp. 331–342). C. E Zsambok and G. Klein, Eds. Lawrence Erlbaum Associates, Inc., Mahwah, NJ

Van Sickle, G. A. 2000. "Situational Awareness From Aircraft Self Reports." [CD-ROM]. In *Symposium Proceedings, Threats, Countermeasures, and Situational Awareness: Teaming for Survivability Symposium and Exhibition.* Electronic Warfare Advanced Technology Program, Virginia Beach, VA. June.

Velleman, P. F. and L. Wilkinson. 1993. "Nominal, Ordinal, Interval, and Ratio Typologies Are Misleading," *The American Statistician, 47*, 65–72.

Vicente, K. J. and J. Rasmussen. 1992. "Ecological Interface Design: Theoretical Foundations," *IEEE Transactions on Systems, Man, and Cybernetics, 22*, 589–605.

Zachary, W. W., A. L. Zaklad, J. H. Hicinbothom, J. M. Ryder, J. A. Purcell, and J. R. Wherry, Jr. 1992. "COGNET Representation of Tactical Decision-Making in Ship-Based Anti-Air Warfare." Tech. Rpt. 920211.9009. CHI Systems, Inc., Spring House, PA.

Zimmerman, L. R. 1997. "Improving Air Combat Identification." Technical Report. Techmatics, Inc., Arlington, VA.

Zsambok, C., L. R. Beach, and G. A. Klein. 1992. "*A Literature Review of Analytical and Naturalistic Decision making* (Tech. Rpt). Naval Command, Control and Ocean Surveillance Center, San Diego, CA. Fairborn, OH: Klein Associates Inc.

Zsambok, C. E. and G. A. Klein. (Eds.). 1997. *Naturalistic Decision Making.* Lawrence Erlbaum Associates, Inc., Mahwah, NJ.

APPENDIX A

CIC ROLE DESCRIPTIONS AND GLOSSARY

The following descriptions are from many sources, including interviews with subject matter experts and definitions provided in *Joint Tactical Terminal (JTT) Operational Analysis Report*[1] and *COGNET Representation of Tactical Decision-Making in Ship-Based Anti-Air Warfare* (Zachary et al., 1992).

Table A-1. CIC role descriptions.

Term	Description
ACS	Air Control Supervisor – Coordinates AICs
AD	Air Defense (formerly Air Warfare)
AIC	Air Intercept Controller – Controls intercept and strike aircraft
AOR	Area of Operational Responsibility – Location in which the ship is operating
ARC	Auxiliary Radar Console Operator, Air Radar Controller, Air Resource Coordinator
ASM	Antiship Missile
ASUWC	Anti-Surface Warfare Coordinator - Responsible for the surface warfare operations.
AW	see AD
AWC/AAWC	Anti-Air Warfare Coordinator – Controls the air defense team, Directs and integrates track assessment, monitoring, and engagement
CIC	Combat Information Center
CAP	Combat Air Patrol
CICWO	CIC Watch Officer
CO	Commanding Officer – Responsible for the overall operation of the ship.
CSC	Combat System Coordinator - Responsible for the operation of the AEGIS combat systems.
CTs	Communications Technicians
ELNOT	ELINT Notation
ES/M	Electronic Support/Measures – Intercept and analysis of electronic emissions
EWO/CO	Electronic Warfare Operator / Console Operator - Maintains the sensor based EW situation.
EWS	Electronic Warfare Supervisor – Control ES and EW systems, Directs EWCOs
FAWC	Force Air Warfare Coordinator
IDS	Identification Supervisor – Coordinates track ID process, Controls IFF system
IFF	Identification Friend or Foe
INTEL	Intelligence Reports
MSS	Missile System Supervisor - Controls the operation of all shipboard missiles.

---

[1] Lyons, *op. cit.*

Table A-1. CIC role descriptions (continued).

Term	Description
OTC	Officer in Tactical Command
RSC	RADAR System Controller
TAO	Tactical Action Officer – Directs and coordinates responses to threats with all warfare coordinators.
TIC	Tactical Information Coordinator – Controls sensor operators and data links
VLS	Vertical Launch Supervisor
XW	See FAWC

APPENDIX B

CUE DEFINITIONS

Table B-1. Assessment cues.

Cue	Description
Airlane	A published or otherwise known commercial air route.
Altitude	Approximate feet above ground or indication of change (e.g., climbing).
Coordinated activity	Track is communicating with, or nearby, another track.
Course	Heading - Exact compass heading or indication of heading relative to own ship (i.e., opening or closing).
CPA	Closest Point of Approach - Estimated distance that track will pass by own ship if track and own ship remain on their current courses.
ES/Radar	Electronic Support - Electronic emissions from the track (typically indicates the type of radar system the track is using).
Feet Wet/Dry	*Feet Dry* track is flying over land. *Feet Wet* track is flying over water.
IFF Mode	Identify friend or foe. Signals from a track that indicates if it is a friendly, or perhaps neutral, aircraft.
Maneuvers	Indicates number of recent maneuvers or if track is following ship.
Number/Composition	Number of aircraft in the formation.
Origin/Location	Indicates the country from which the track most likely originated.
Own Support	Availability of nearby friendly ships or patrol aircraft (CAP)
Range/Distance	The track's distance from own ship.
Speed	Approximate airspeed or an indication of change (e.g., increasing).
Visibility	Approximate number of miles, or an indication of atmospheric conditions (e.g., haze).
Weapon envelope	Track's position to its estimated weapons envelope (e.g., within).
Wings Clean/Dirty	A track without weapons is designated as *Wings Clean*. A track with weapons is designated *Wings Dirty*. Determined by visual identifaction by intercepting aircraft.
Confidence in equipment	Degree of confidence in the equipment (Combat Direction System).
Confidence in Team	Degree of confidence in the Air Warfare team.
Voice Communication	Track has or has not responded to inquiries or warnings.
Weapons	Type or category (e.g., air to air) of weapon carried by the track.
Weapons Status (Own)	Alert status of the ship and the friendly aircraft under its control.

APPENDIX C

STUDY 3.1 CUE QUESTIONNAIRE

The questionnaire was reformatted to fit within the margins of this report.

AIRCRAFT TRACK ANALYSIS

Purpose

- The purpose of this questionnaire is to gather specific information about the data that are used by the air defense team when they analyze radar contacts.

- Many of the questions are about single pieces of information (e.g., altitude). We realize that information is not processed in isolation in the real situation. However, we need to understand how individual elements are handled.

- The next page covers some of your background.

BACKGROUND INFORMATION

Indicate the information to the best of your recollection.

Current Rank/Rate: _____   Time: _____years

Air Warfare/Defense Experience within the Combat Information Center (CIC):

    <u>Please check all that apply</u>   <u>Total at-sea time</u>

_____ Commanding Officer (CO) .... _____years

_____ Tactical Action Officer (TAO).......................... _____years

_____ CIC Watch Officer .... _____years

_____ AntiAir Warfare Coordinator (AAWC) .... _____years

_____ Tactical Information Coordinator (TIC) .... _____years

_____ Air Resource Coordinator (ARC) .... _____years

_____ Identification Supervisor (IDS)........................ _____years

_____ Air Control Supervisor (ACS)......................... _____years

_____ Air Coordinator (AIC)..................................... _____years

_____ Electronic Warfare Supervisor (EWS)............. _____years

_____ Electronic Warfare Supervisor (EWS)............. _____years

_____ EW Console Operator (EWCO)....................... _____years

_____ Missile Sysytem Supervisor (MSS) .... _____years

_____ Vertical Launch Supervisor (VLS) .... _____years

_____ Other Air Warfare/Defense Experience .... _____years

Please describe other AW experience(s):

_____

_____

Qualifications:

    Surface Warfare Officer (1110, etc.) qualified…………………….._____yes _____no

    Enlisted Surface Warfare Specialist qualified…………………..._____yes _____no

Please STOP here and wait for instructions.

INSTRUCTIONS:

- Treat each item (e.g., speed) as if were the first piece of data you received.
- Indicate how the data would CHANGE your estimate of threat.
- Each question requires two answers - Answer as if you were in the following Areas of Responsibility:
  - o Littoral AOR (e.g., Northern Persian Gulf)
  - o Open Ocean AOR (e.g., Southern Pacific)
- Please answer the questions as accurately as possible.
- We are interested in your <u>at-sea</u> experience, in what has worked for you.

1. TYPES OF AIRCRAFT – How threatening, on average, is this type of aircraft, based on your experience?

	"Littoral" AOR (e.g., Persian Gulf)					"Open" AOR (e.g., Southern Pacific)				
	never	rarely	sometimes	often	always	never	rarely	sometimes	often	always
Friend	1	2	3	4	5	1	2	3	4	5
Unknown, Assumed Friend	1	2	3	4	5	1	2	3	4	5
Unknown	1	2	3	4	5	1	2	3	4	5
Unknown, Assumed Enemy	1	2	3	4	5	1	2	3	4	5
Hostile	1	2	3	4	5	1	2	3	4	5

2. WEAPONS - What is the level of threat posed to your ship/battle group by this type of weapon on an Unknown Aircraft?

	"Littoral" AOR (e.g., Persian Gulf)					"Open" AOR (e.g., Southern Pacific)				
	never	rarely	sometimes	often	always	never	rarely	sometimes	often	always
Missile										
Exocet	1	2	3	4	5	1	2	3	4	5
Harpoon	1	2	3	4	5	1	2	3	4	5
Maverick	1	2	3	4	5	1	2	3	4	5
Other _____	1	2	3	4	5	1	2	3	4	5
Other _____	1	2	3	4	5	1	2	3	4	5
Guns	1	2	3	4	5	1	2	3	4	5
Bombs	1	2	3	4	5	1	2	3	4	5
Other _____	1	2	3	4	5	1	2	3	4	5

3. RADAR (ES) EMITTERS - Please rate, according to your experience, by checking the description that best fits in the blank.

This emitter is _____ associated with potentially threatening aircraft.

APG-63.........____ always	____often	____sometimes	____hardly ever	____never
APS-24.........____ always	____often	____sometimes	____hardly ever	____never
APS-115........____ always	____often	____sometimes	____hardly ever	____never
ARINC-564....____ always	____often	____sometimes	____hardly ever	____never
CAS Search....____ always	____often	____sometimes	____hardly ever	____never
Castor II.........____ always	____often	____sometimes	____hardly ever	____never
Cyrano IV......____ always	____often	____sometimes	____hardly ever	____never
Decca-170......____always	____often	____sometimes	____hardly ever	____never
Decca-1226.....____always	____often	____sometimes	____hardly ever	____never
Don-2............____ always	____often	____sometimes	____hardly ever	____never
Heracles-2......____always	____often	____sometimes	____hardly ever	____never
LN-66............____always	____often	____sometimes	____hardly ever	____never
Primus-40.......____always	____often	____sometimes	____hardly ever	____never
RDR-1500......____always	____often	____sometimes	____hardly ever	____never
SPS-10B....... ____always	____ often	____sometimes	____hardly ever	____never
SPS-55...........____always	____often	____sometimes	____hardly ever	____never
SPS-49...........____always	____often	____sometimes	____hardly ever	____never
Type-992Q......____always	____often	____sometimes	____hardly ever	____never
Type-1006......____always	____often	____sometimes	____hardly ever	____never
Triton............____always	____often	____sometimes	____hardly ever	____never
Other: ....____always	____often	____sometimes	____hardly ever	____never

_____

FLIGHT PARAMETERS - How would the values in the ranges below CHANGE your estimate of threat?

Values in this range	"Littoral" AOR (e.g., Persian Gulf)					"Open" AOR (e.g., Southern Pacific)				
	lower greatly	lower a little	no change	raise a little	raise greatly	lower greatly	lower a little	no change	raise a little	raise greatly
4. Speed										
Speed steady	1	2	3	4	5	1	2	3	4	5
Speed increase	1	2	3	4	5	1	2	3	4	5
Speed decrease	1	2	3	4	5	1	2	3	4	5
Speed of										
under 150 kts	1	2	3	4	5	1	2	3	4	5
150-250 kts	1	2	3	4	5	1	2	3	4	5
250-350 kts	1	2	3	4	5	1	2	3	4	5
350-450 kts	1	2	3	4	5	1	2	3	4	5
450-550 kts	1	2	3	4	5	1	2	3	4	5
over 550 kts	1	2	3	4	5	1	2	3	4	5
5. Altitude										
Altitude steady	1	2	3	4	5	1	2	3	4	5
Altitude increase	1	2	3	4	5	1	2	3	4	5
Altitude decrease	1	2	3	4	5	1	2	3	4	5
Altitude of										
under 500 ft	1	2	3	4	5	1	2	3	4	5
500-1000 ft	1	2	3	4	5	1	2	3	4	5
1000-5000 ft	1	2	3	4	5	1	2	3	4	5
5000-10000 ft	1	2	3	4	5	1	2	3	4	5
10000-20000 ft	1	2	3	4	5	1	2	3	4	5
over 20000 ft	1	2	3	4	5	1	2	3	4	5
6. Course										
Steady and closing	1	2	3	4	5	1	2	3	4	5
Steady and opening	1	2	3	4	5	1	2	3	4	5
Turn to closing	1	2	3	4	5	1	2	3	4	5
Turn to opening	1	2	3	4	5	1	2	3	4	5
CPA under 1 nm	1	2	3	4	5	1	2	3	4	5
1-5 nm	1	2	3	4	5	1	2	3	4	5
5-15 nm	1	2	3	4	5	1	2	3	4	5
15-25 nm	1	2	3	4	5	1	2	3	4	5
25-50 nm	1	2	3	4	5	1	2	3	4	5
over 50 nm	1	2	3	4	5	1	2	3	4	5
Many maneuvers	1	2	3	4	5	1	2	3	4	5
Track's response to own-ship movements										
changes course to match	1	2	3	4	5	1	2	3	4	5
remains on course	1	2	3	4	5	1	2	3	4	5

Values in this range	"Littoral" AOR (e.g., Persian Gulf)					"Open" AOR (e.g., Southern Pacific)				
	lower greatly	lower a little	no change	raise a little	raise greatly	lower greatly	lower a little	no change	raise a little	raise greatly
7. Distance from Own Ship										
under 5 nm	1	2	3	4	5	1	2	3	4	5
5-25 nm	1	2	3	4	5	1	2	3	4	5
25-50 nm	1	2	3	4	5	1	2	3	4	5
over 50 nm	1	2	3	4	5	1	2	3	4	5
8. Weapons										
wings dirty	1	2	3	4	5	1	2	3	4	5
wings clean	1	2	3	4	5	1	2	3	4	5
capability unknown	1	2	3	4	5	1	2	3	4	5
9. Weapon Envelope - When track is…										
too close to shoot	1	2	3	4	5	1	2	3	4	5
within (near inner edge)	1	2	3	4	5	1	2	3	4	5
within	1	2	3	4	5	1	2	3	4	5
within (near outer edge)	1	2	3	4	5	1	2	3	4	5
outside (too far away)	1	2	3	4	5	1	2	3	4	5
10. IFF Mode										
none	1	2	3	4	5	1	2	3	4	5
1	1	2	3	4	5	1	2	3	4	5
2	1	2	3	4	5	1	2	3	4	5
3	1	2	3	4	5	1	2	3	4	5
4	1	2	3	4	5	1	2	3	4	5
1, 2	1	2	3	4	5	1	2	3	4	5
2, 3	1	2	3	4	5	1	2	3	4	5
1, 2, 3	1	2	3	4	5	1	2	3	4	5
C	1	2	3	4	5	1	2	3	4	5
11. Number of Aircraft										
unknown	1	2	3	4	5	1	2	3	4	5
single contact	1	2	3	4	5	1	2	3	4	5
3 - 5 in group	1	2	3	4	5	1	2	3	4	5
over 5 in group	1	2	3	4	5	1	2	3	4	5

Values in this range	"Littoral" AOR (e.g., Persian Gulf)					"Open" AOR (e.g., Southern Pacific)				
	lower greatly	lower a little	no change	raise a little	raise greatly	lower greatly	lower a little	no change	raise a little	raise greatly

12. Airlane

Values in this range	L greatly	L little	no chg	R little	R greatly	O greatly	O little	no chg	R little	R greatly
on/following	1	2	3	4	5	1	2	3	4	5
parallel and										
under 5 nm off	1	2	3	4	5	1	2	3	4	5
5 - 10 nm off	1	2	3	4	5	1	2	3	4	5
over 10 nm off	1	2	3	4	5	1	2	3	4	5
perpendicular and										
under 5 nm off	1	2	3	4	5	1	2	3	4	5
5 - 10 nm off	1	2	3	4	5	1	2	3	4	5
over 10 nm off	1	2	3	4	5	1	2	3	4	5

13. Radar

Values in this range	L greatly	L little	no chg	R little	R greatly	O greatly	O little	no chg	R little	R greatly
constant radar return	1	2	3	4	5	1	2	3	4	5
intermittent radar return	1	2	3	4	5	1	2	3	4	5
no emitter	1	2	3	4	5	1	2	3	4	5
illuminates own ship	1	2	3	4	5	1	2	3	4	5
possibly providing targeting	1	2	3	4	5	1	2	3	4	5

14. Feet Wet/Dry

Values in this range	L greatly	L little	no chg	R little	R greatly	O greatly	O little	no chg	R little	R greatly
dry	1	2	3	4	5	1	2	3	4	5
wet, parallel to coast and										
under 5 nm from coast	1	2	3	4	5	1	2	3	4	5
over 5 nm from coast	1	2	3	4	5	1	2	3	4	5
wet, flying away from coast										
under 5 nm from coast	1	2	3	4	5	1	2	3	4	5
over 5 nm from coast	1	2	3	4	5	1	2	3	4	5

15. Origin/Location

Values in this range	L greatly	L little	no chg	R little	R greatly	O greatly	O little	no chg	R little	R greatly
known friend	1	2	3	4	5	1	2	3	4	5
known hostile	1	2	3	4	5	1	2	3	4	5
known neutral	1	2	3	4	5	1	2	3	4	5
route suggests friend	1	2	3	4	5	1	2	3	4	5
route suggests hostile	1	2	3	4	5	1	2	3	4	5
route suggests neutral	1	2	3	4	5	1	2	3	4	5

16. Voice communication with track

Values in this range	L greatly	L little	no chg	R little	R greatly	O greatly	O little	no chg	R little	R greatly
none	1	2	3	4	5	1	2	3	4	5
have comm.	1	2	3	4	5	1	2	3	4	5

17. Own support in area

Values in this range	L greatly	L little	no chg	R little	R greatly	O greatly	O little	no chg	R little	R greatly
none available	1	2	3	4	5	1	2	3	4	5
surface available	1	2	3	4	5	1	2	3	4	5
air (DCA) available	1	2	3	4	5	1	2	3	4	5

	"Littoral" AOR (e.g., Persian Gulf)					"Open" AOR (e.g., Southern Pacific)				
Values in this range	lower greatly	lower a little	no change	raise a little	raise greatly	lower greatly	lower a little	no change	raise a little	raise greatly
18. Warning status										
white	1	2	3	4	5	1	2	3	4	5
yellow	1	2	3	4	5	1	2	3	4	5
red	1	2	3	4	5	1	2	3	4	5
19. Confidence in data/equipment										
High	1	2	3	4	5	1	2	3	4	5
Low	1	2	3	4	5	1	2	3	4	5
20. Confidence in fellow team members										
High	1	2	3	4	5	1	2	3	4	5
Low	1	2	3	4	5	1	2	3	4	5

21. Intelligence Reports and Briefs

List the top 3 types of information that would affect aircraft assessment (examples: terrorist activity, troop/ship movements, and other geopolitical news):

Effect on Threat Estimate

	"Littoral" AOR (e.g., Persian Gulf)					"Open" AOR (e.g., Southern Pacific)				
Information	lower greatly	lower a little	no change	raise a little	raise greatly	lower greatly	lower a little	no change	raise a little	raise greatly
1. _____ _____.........	1	2	3	4	5	1	2	3	4	5
2. _____ _____.........	1	2	3	4	5	1	2	3	4	5
3. _____ _____.........	1	2	3	4	5	1	2	3	4	5

22. Air Defense Information Flow

Write your position in the box below.

On the arrows, write the information you NORMALLY receive and pass on.

You may draw more lines if needed.

List the positions and data that you usually GIVE information to:

_____ _____ _____ _____ _____ _____

```
        ↖      ↖    ↑    ↑    ↗        ↗
```

┌──────────────────────┐
│ **Your Position:**   │
│   │
│   │
└──────────────────────┘

```
        ↗    ↗   ↑   ↖   ↖
```

_____ _____ _____ _____ _____ _____
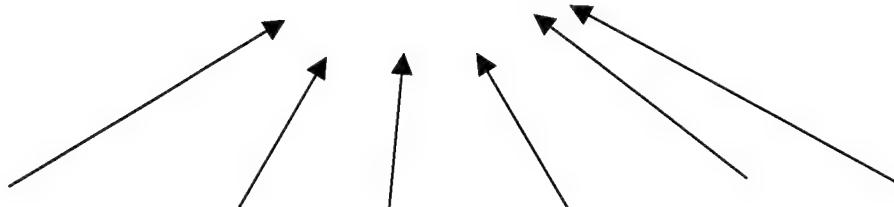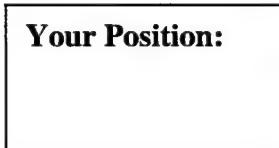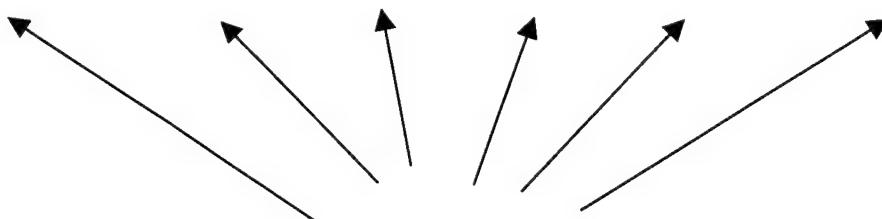
List the positions and data that you usually RECEIVE information from

23. Air Defense Information

List the information/data that you normally use to evaluate an aircraft track, in order of importance to you:

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____

24. Combination of Information

What information do you consider in combination with other information (for example, speed & range)?

Please list your **most often used combinations:**

_____

_____

_____

_____

_____

_____

THANK YOU. You are finished!

APPENDIX D

STUDY 1 QUESTIONNAIRE RESULTS

Participants responded to all questions on the questionnaire, however, only selected responses are reported here. Contact the authors for additional data.

Table D-1. Mean Threat Level for each radar emitter.

Radar Emitter	Threat Level
rdr1500	3.0

Table D-2. Threat Change Ratings for each cue.

Cue	Data Value or Range	Littoral	Open Ocean
Airlane	On airlane	-1.0	-1.0
	Parallel, < 5 nmi off	-0.6	-1.0
Altitude	Steady	-0.2	-0.5
	Increasing	0.0	-0.5
Confidence in equipment			
	High	0.0	-0.5
	Low	1.2	0.5
Confidence in team			
	High	0.2	-0.5
	Low	1.2	1.5
Coordinated activity	Targeting	2.0	2.0
Course	Steady and Closing	1.8	1.0
CPA	25 to 50 nmi	0.2	-1.5
	Over 50 nmi	-1.0	-1.5
ES/Radar	Constant return	0.0	0.0
Feet wet	Feet dry	-0.4	-1.5
IFF mode	None	1.6	2.0
	3	0.2	0.0
Maneuvers	Many maneuvers	1.2	1.5
Number	Unknown	1.5	1.5
Origin	known friendly	-0.8	-2.0
Own Support	CAP aircraft	0.0	-0.5
Range	25 to 50 nmi	0.8	0.0
	Over 50 nmi	-0.4	-1.0
Speed	Steady	0.0	0.0
Visibility	Good	- 0.3	0.0
Voice communications	Yes, response to warning	-0.4	-0.5
Weapons			
Weapon envelope	Within	1.6	1.0
Weapons status (Own)	White	0.0	-0.5
Wings	Wings unknown	1.2	1.0

APPENDIX E

STUDY 3 DESCRIPTIVE STATISTICS

Table E-1. Mean ID assignment (Neutral, Unknown, or Suspect) for each combination of ID-Platform and Position of Conflicting Data.

Given ID-Platform Position of Conflicting Data	Neutral COMAIR			Neutral MARPAT			Suspect TACAIR			
	Mean	Std. Dev.	Approx. ID	Mean	Std. Dev.	Approx. ID	Mean	Std. Dev.	Approx. ID	N
None	1.27	0.46	Neutral	1.27	0.59	Neutral	2.87	0.52	Suspect	15
Low	1.23	0.57	Neutral	1.63	0.76	Unknown	2.63	0.67	Suspect	30
Medium	1.47	0.73	Neutral	1.40	0.67	Neutral	2.80	0.55	Suspect	30
High	1.97	0.81	Unknown	2.30	0.79	Unknown	2.00	0.87	Unknown	30
Low+high	2.27	0.83	Unknown	2.10	0.80	Unknown	2.00	0.87	Unknown	30
Low+medium	2.03	0.89	Unknown	1.67	0.71	Unknown	2.50	0.73	Suspect	30
Medium+high	2.27	0.78	Unknown	2.33	0.88	Unknown	2.40	0.77	Unknown	30
All	2.03	0.85	Unknown	1.77	0.94	Unknown	2.10	0.92	Unknown	30
Total	1.85	0.85		1.84	0.85	Unknown	2.38	0.82	Unknown	225

Table E-2. Mean Time (in seconds) to assign an ID for each combination of ID-Platform and Position of Conflicting Data.

Given ID-Platform Position of Conflicting Data	Neutral COMAIR		Neutral MARPAT		Suspect TACAIR		
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	N
None	26119.13	26672.99	16649.13	10587.31	18473.60	11596.68	15
Low	23750.97	21934.37	25776.47	17262.60	19168.47	11682.83	30
Medium	26676.93	23915.12	17543.73	12183.33	21828.47	17445.51	30
High	25834.07	34071.95	29602.97	46804.99	23697.87	15958.54	30
Low+high	18581.00	12977.35	26797.03	15877.10	30654.60	28097.12	30
Low+medium	22887.30	13642.54	18735.83	13377.86	28229.53	17108.19	30
Medium+high	23942.53	15267.35	22629.83	21394.56	24437.50	16027.36	30
All	24048.37	11751.56	27996.83	43639.63	44424.77	138409.87	30
Total	23837.43	20756.68	23654.30	27035.12	26890.40	53043.72	225

Table E-3. Mean Platform assignment (COMAIR, HELO, MARPAT, or TACAIR) for each combination of ID-Platform and Position of Conflicting Data.

Given ID-Platform	Neutral COMAIR			Neutral MARPAT			Suspect TACAIR			
Position of Conflicting Data	Mean	Std. Dev.	Approx. Platform	Mean	Std. Dev.	Approx. Platform	Mean	Std. Dev.	Approx. Platform	N
None	1.20	0.77	COMAIR	2.87	0.52	MARPAT	4.00	0.00	TACAIR	15
Low	2.57	1.28	MARPAT	3.17	0.75	MARPAT	3.37	0.72	MARPAT	30
Medium	1.47	1.01	COMAIR	2.93	0.58	MARPAT	3.80	0.61	TACAIR	30
High	1.50	0.94	MARPAT	3.03	0.18	MARPAT	3.80	0.41	TACAIR	30
Low+high	2.80	1.27	MARPAT	3.23	0.43	MARPAT	2.67	1.27	MARPAT	30
Low+medium	3.13	0.97	MARPAT	3.03	0.93	MARPAT	2.77	1.19	MARPAT	30
Medium+high	2.10	1.12	HELO	2.90	0.71	MARPAT	3.80	0.61	TACAIR	30
All	2.67	1.18	MARPAT	2.63	1.13	MARPAT	2.83	1.44	TACAIR	30
Total	2.24	1.26		2.98	0.73		3.34	1.05		225

Table E-4. Mean Time (in seconds) to assign a Platform for each combination of ID-Platform and Position of Conflicting Data.

Given ID-Platform	Neutral COMAIR		Neutral MARPAT		Suspect TACAIR		
Position of Conflicting Data	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	N
None	1485.73	1163.23	2410.80	2120.61	1737.73	1924.35	15
Low	6542.53	14404.48	4239.60	5325.52	4151.93	5600.31	30
Medium	5807.07	7747.31	2996.90	3085.79	2897.43	4489.57	30
High	3636.60	6227.68	2884.83	5626.08	4301.40	5352.82	30
Low+high	3767.00	3886.28	5800.20	9376.53	3907.03	4425.25	30
Low+medium	3804.23	6797.82	6276.93	12639.60	6687.10	16799.80	30
Medium+high	5793.00	7267.54	2490.20	2487.71	3037.73	3281.83	30
All	7257.97	11514.61	2720.60	3045.29	2730.00	1924.35	15
Total	4980.17	8621.10	3815.29	6741.49	3810.87	5600.31	30

Table E-5. Mean Threat Level assignment (1=Low to 7=High) for each combination of Position of Conflicting Data and ID-Platform.

Given ID-Platform	Neutral COMAIR		Neutral MARPAT		Suspect TACAIR		
Position of Conflicting Data	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	N
None	0.93	1.03	1.13	1.30	5.07	0.96	15
Medium	1.37	1.22	1.60	1.25	3.77	1.41	30
Low	1.50	1.50	2.27	1.46	3.33	1.32	30
High	1.80	1.40	3.10	1.27	2.77	1.22	30
Low+high	2.67	1.42	3.20	1.19	2.47	1.41	30
Low+medium	3.10	1.63	1.73	1.23	2.87	1.20	30
Medium+high	2.57	1.30	3.10	1.45	3.37	1.07	30
All	2.33	1.15	2.33	1.83	2.83	1.80	30
Total	2.11	1.50	2.39	1.53	3.19	1.47	225

Table E-6. Mean Time (seconds) to assign a Threat Level for each combination of Position of Conflicting Data and ID-Platform.

Given ID-Platform	Neutral COMAIR		Neutral MARPAT		Suspect TACAIR		
Position of Conflicting Data	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	N
None	3843.80	4637.83	3656.80	2809.57	6078.33	8244.78	15
Low	4810.80	7883.91	3894.90	2823.77	4391.17	3667.68	30
Medium	2339.67	1561.64	3228.23	4496.43	6358.23	8690.79	30
High	3424.47	3110.54	5350.63	5964.77	10782.33	29219.69	30
Low+high	3572.17	3597.66	4757.60	4033.72	4031.63	4347.68	30
Low+medium	4129.73	4259.80	3371.63	2371.20	4437.67	4365.41	30
Medium+high	5180.43	3839.99	3825.47	3774.66	4692.93	4108.80	30
All	4014.80	2916.47	3458.50	3358.52	9125.00	28649.44	30
Total	3919.20	4324.92	3962.05	3923.49	6247.75	15665.73	225

Table E-7. Mean Intent assignment (1=Friendly to 7=Hostile) for each combination of ID-Platform and Position of Conflicting Data.

Given ID-Platform	Neutral COMAIR			Neutral MARPAT			Suspect TACAIR			
Position of Conflicting Data	Mean	Std. Dev.	Approx. Intent	Mean	Std. Dev.	Approx. Intent	Mean	Std. Dev.	Approx. Intent	N
None	0.87	0.99	Friendly	1.27	1.28	Friendly	4.47	0.74		15
Low	1.50	1.28	Friendly	1.87	1.20	Probably friend	3.13	1.31	Probably friendly	30
Medium	1.27	1.14	Friendly	1.57	1.10	Probably friend	3.83	1.51	Uncertain	30
High	2.03	1.61	Friendly	3.23	1.28	Probably friendly	2.57	1.01	Probably friendly	30
Low+high	2.47	1.38	Probably friendly	3.17	1.23	Probably friendly	2.43	1.28	Probably friendly	30
Low+medium	2.90	1.42	Probably friendly	1.80	1.37	Probably friendly	2.90	1.27	Probably friendly	30
Medium+high	2.70	1.26	Probably friendly	3.23	1.48	Probably friendly	2.87	1.01	Probably friendly	30
All	2.50	1.01	Probably friendly	2.40	1.79	Probably friendly	2.47	1.43	Probably friendly	30
Total	2.11	1.43		2.39	1.53		2.99	1.36		225

E-8. Mean Confidence Level assignment (1=Low to 7=High) for each combination of ID-Platform and Position of Conflicting Data.

Given ID-Platform	Neutral COMAIR		Neutral MARPAT		Suspect TACAIR		
Position of Conflicting Data	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	N
None	4.67	1.59	5.00	1.00	4.93	1.03	15
Low	4.63	1.30	4.53	1.17	4.47	1.33	30
Medium	4.23	1.72	4.43	1.30	4.93	1.14	30
High	4.57	1.30	4.73	1.34	4.27	1.11	30
Low+high	4.70	1.34	4.57	1.14	4.33	1.30	30
Low+medium	4.27	1.48	4.43	1.19	4.30	1.21	30
Medium+high	4.23	1.25	4.90	1.03	4.40	1.33	30
All	4.33	1.52	4.63	1.03	4.73	1.08	30
Total	4.44	1.43	4.63	1.16	4.52	1.21	225

Table E-9. Mean Time (in seconds) to assign a Confidence Level for each combination of ID-Platform and Position of Conflicting Data.

Given ID-Platform	Neutral COMAIR		Neutral MARPAT		Suspect TACAIR		
Position of Conflicting Data	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	N
None	1754.40	756.21	1987.60	1213.30	1521.60	1198.98	15
Low	1361.23	688.41	1867.40	1442.28	1339.00	706.22	30
Medium	1472.97	958.54	1366.53	1057.38	1684.07	1724.69	30
High	1634.60	1742.18	1900.73	1815.95	1631.67	1142.86	30
Low+high	1415.33	734.23	1919.27	1799.57	1731.83	1187.19	30
Low+medium	1444.33	803.75	1471.20	964.49	1486.83	1031.54	30
Medium+high	1837.53	1225.51	1825.77	1745.81	1780.37	1823.36	30
All	1578.10	942.33	2467.23	2142.32	2179.33	2082.41	30
Total	1549.51	1047.02	1841.59	1601.66	1679.19	1441.65	225

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
08–2002	Final	

4. TITLE AND SUBTITLE

STUDIES OF U.S. NAVY CUES, INFORMATION ORDER, AND IMPACT OF CONFLICTING DATA

5a. CONTRACT NUMBER
N66001-99-D-0050

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
0602235N

6. AUTHORS

M. J. Liebhaber
D. A. Kobus
Pacific Science and Engineering Group, Inc.
6310 Greenwich Drive, Suite 200
San Diego, CA 92122

B. A. Feher
SSC San Diego

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

SSC San Diego
San Diego, CA 92152–5001

8. PERFORMING ORGANIZATION REPORT NUMBER
TR 1888

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research, (PMA-282)
Human Systems Department
800 North Quincy Street
Arlington, VA 22217–5660

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

This is the work of the United States Government and therefore is not copyrighted. This work may be copied and disseminated without restriction. Many SSC San Diego public release documents are available in electronic format at:
http://www.spawar.navy.mil/sti/publications/pubs/index.html

14. ABSTRACT

This report discusses the results of three studies investigating the threat assessment process for shipboard air defense (AD). The goal was to better understand the relationship between air track information (e.g., altitude, speed, country of origin) available to a ship's AD personnel and their perceived level of threat regarding a particular aircraft. Understanding this process is crucial in designing effective AD decision support tools.

15. SUBJECT TERMS

Mission Area: Command and Control
threat assessment
human–computer interface
decision support system

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			B. A. Feher, Code 244210
U	U	U	UU	90	**19B. TELEPHONE NUMBER** *(Include area code)* (619) 553–9226

INITIAL DISTRIBUTION

20012	Patent Counsel	(1)
20271	Archive/Stock	(6)
20274	Library	(2)
2027	M. E. Cathcart	(1)
20271	E. R. Ratliff	(1)
20271	D. Richter	(1)
244210	B. A. Feher	(9)

Defense Technical Information Center
Fort Belvoir, VA 22060–6218 (4)

SSC San Diego Liaison Office
C/O PEO-SCS
Arlington, VA 22202–4804

Center for Naval Analyses
Alexandria, VA 22311–1850

Office of Naval Research
ATTN: NARDIC (Code 362)
Arlington, VA 22217–5660

Government-Industry Data Exchange
    Program Operations Center
Corona, CA 91718–8000